Towards Structural Understanding of Phospholipase PlaB from *Legionella pneumophila*

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades einer

Doktorin der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Pham, Thanh Van

aus Hanoi / Vietnam

1. Referentin oder Referent:	Professor Dr. Dirk W. Heinz
2. Referentin oder Referent:	Professor Dr. Dieter Jahn
eingereicht am:	28.10.2014
mündliche Prüfung (Disputation) am:	28.01.2015

Druckjahr 2015

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Tagungsbeiträge

Pham, S.T.V, Kuhle, K., Scrima, A., Kraußе, J., Flieger, A. & Heinz, D.W.: Toward a Structural Understanding of the Activities of Phospholipase PlaB, a Novel Virulence Factor from *Legionella pneumophila. Oral presentation at 15th Heart of Europe Biocrystallography Meeting,* Beilngries, Germany (2012).

Contents

Abbreviations

Å	Ångström
A280	UV absorbance at wavelength of 280 nm
Amp	Ampicillin
Community-acquired pneumonia	CAP
DLS	Dynamic light scattering
DMSO	Dimethyl sulfoxide
DPPC	Dipalmitoylphosphatidylcholine
DTT	Dithiothreitol
DVDC	Drop vapor diffusion crystallization
E. coli	*Escherichia coli*
ERSF	European Synchrotron Radiation Facility
ESI-MS	Electrospray Ionisation
HABA	Hydroxy-azophenyl-benzoic acid
HZI	Helmholtz-Zentrum für Infektionsforschung
IPTG	Isopropyl-β-D-thiogalactopyranoside
L. pneumophila	*Legionella pneumophila*
LPC	Lysophosphatidylcholine
LPLA	Lysophospholipase A
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation – Time Of Flight
MR	Molecular replacement
MS	Mass spectrometry
MWCO	Molecular weight cutoff
OD_{600}	Optical density at wavelength of 600 nm
OG	Octyl-β-D-glucopyranoside
PC	Phosphatidylcholine
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PG	Phosphatidylglycerol
PKCα	Protein kinase Cα
PLA	Phospholipase A
R_H	Hydrodynamic radius
RKI	Robert Koch Institute
Rpm	Rotations per minute
R.T.	Room temperature
SD200	Superdex 200 (gel filtration column)
SDS	Sodium dodecyl sulfate

SDS-PAGE	SDS polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SBDD	Structure-based drug design
SeMet	Selenium methionine

Zusammenfassung

Legionella pneumophila (*L. pneumophila*) ist die Hauptursache für Legionellose, deren ernste Form, die Legionärskrankheit, tödlich verlaufen kann. In letzter Zeit wurde in Deutschland ein starker Anstieg der Legionellose als Ursache für ambulant erworbene Pneumonie sowohl in stationär als auch in ambulant behandelten Krankenhauspatienten beobachtet. Die Bakterien besitzen eine Vielzahl von Phospholipasen, welche Phospholipide spalten und dadurch verschiedenen Produkte freisetzen, die die bakterielle Invasion und Virulenz unterstützen. Zu ihnen gehört PlaB, die vorherrschende zell-assoziierte Phospholipase deren Aktivität zu Lungenentzündung und markantem Velust von funktionaler epithelialer Zellschicht und somit Zersörung von Lungengewebe führt. Der detailierte Mechanismus der PlaB-Aktivität ist noch unbekannt. Es ist darüber hinaus ein neuartiges Protein für welches noch keine homologe Struktur zur Verfügung steht. Strukturelle Information über PlaB könnte deshalb die Grundlage bilden für die Aufklärung des enzymatischen Mechanismus im Allgemeinen und der PlaB-Familie im Besonderen. Die Sruktur des Komplexes aus dem Protein und seinen Substraten würde auch die Möglichkeit zur Entdeckung neuer Medikamente bieten, welche antibakterielle Resistenzen umgehen könnten.

In der vorliegenden Arbeit wurden Wege zur großmaßstäblichen Expression und Reinigung von PlaB und seinen SeMet-markierten Derivaten etabliert, und das Protein wurde bis zur Homogenität gereinigt. Kristalle wurden gezüchtet, und ein nativer Datensatz mit einer Auflösung von 2.6 Å wurde gesammelt. Alle Kristalle waren jedoch entweder verzwillingt, wie im Falle der hexagonalen Kristallform, oder oder sehr stark anisotrop, wie im Falle der triklinen Kristallform. Das machte es unmöglich, das Phasenproblem zu lösen und die Struktur zu bestimmen. Externe Parameter und chemische Modifikationen des Proteins wurden untersucht, um diese Probleme zu lösen. Kristalle mit anderer Morphologie konnten von einer C-terminal verkürzten Version des Proteins gezüchtet werden, was Anlass zur Hoffnung gibt, dass diesen nicht die gleichen Probleme anhaften, die zum Haupthindernis dieser Arbeit wurden.

Summary

Legionella pneumophila (L. pneumophila) is the major cause of Legionellosis, of which the severe form, Legionnaires' disease, can lead to death. Recently, it is increasingly observed as a leading cause of community-acquired pneumonia (CAP) in hospitalized and ambulatory patients in Germany. The bacteria possess a variety of phospholipases that cleave phospholipids, releasing various products which facilitate the invasion and virulence of bacteria. Among them, PlaB is the major cell-associated phospholipase, the activity of which leads to lung inflammation and prominent loss of functional epithelial cell layer in the lung and hence to the destruction of lung tissue. The detailed mechanism of the PlaB activity remains unknown. It is also a novel protein for which no structural homolog is available. Structural information of PlaB may thus establish a foundation to elucidate enzymatic mechanism of the enzyme in particular and PlaB family in general. Structure of a complex between the protein and its substrates will also offer lead for the discovery of novel drug to counteract antibacterial resistance.

In this work, schemes for large-scale expression and purification of PlaB and SeMet-labeled derivatives were established, and the protein was purified to homogeneity. Crystals were grown and a native data set of the protein crystal was obtained at the resolution of 2.6 Å. However, the crystals were either twinned when it was in hexagonal form or seriously anisotropic in triclinic form. This made it impossible to obtain phase information for structural determination. External parameters as well as chemical modifications of the protein itself were investigated intensively to solve these issues. A different crystal morphology of a C-terminal truncated version of the protein was obtained, giving hope that it would not have these two issues that have been the bottleneck of this work.

1 Introduction

1.1 *L. pneumophila* and pneumonia diseases

Gram-negative *L. pneumophila* was first discovered by McDade *et al.* (1977) after an outbreak of Legionnaires' disease, a severe fibrinopurulent pneumonia. The outbreak occurred among veterans at a hotel during an American Legion convention in Philadelphia in 1976, resulting in 29 deaths and the hospitalization of 147 people (Fraser *et al*. 1997). The pathogen was named in memory of the deceased veterans. The air-conditioning system was later identified as the source of the infection during the convention (Kwaik *et al*. 1998). The oldest confirmed outbreak, however, was reported to occur in 1957 at Minnesota meat packing plant (Osterholm *et al*. 1983).

L. pneumophila belongs to the genus *Legionella*, which consists of over 50 species identified to date. Twenty-four of these species have been found to be associated with human disease, but *L. pneumophila* is responsible for approximately 90% of Legionnaires' disease cases (Hilbi *et al*. 2010; Diederen 2008; Fields *et al*. 2002; Muder *et al*. 1986; Yu *et al*. 2002; Marston *et al*. 1997). *L. pneumophila* serogroup 1 is responsible for 84% of cases worldwide caused by *L. pneumophila*, among at least 15 serogroups (Yu *et al*. 2002). Previous studies provided evidences that *L. pneumophila* is more pathogenic to human than other *Legionella* species (Newton *et al*. 2010).

The latest record in 2010 indicated a significant increase in the incidence of *L. pneumophila* diseases in general and Legionnaires' disease in particular. In Germany, the bacterium is a leading cause of CAP in hospitalized and ambulatory patients (Carratalà & Garcia-Vidal 2010). The mortality of severe *L. pneumophila* is high (30%) and poor outcomes are correlated with pre-existing comorbidities (e.g. cardiac disease, diabetes, acute renal failure). Smoking, chronic obstructive lung disease, diabetes, chronic corticosteroid therapy, and weakened immune systems are known risk factors for Legionnaires' disease (Nicolini *et al*. 2013; Hilbi *et al*. 2010).

1.2 State of treatment and drug resistance

Treatment therapy for patients hospitalized with community-acquired pneumonia is established empirically (Fields *et al*. 2002). Previous study showed that there has been an association between delay in starting appropriate therapy and increased mortality (Heath *et al*. 1996). *L. pneumophila* has been shown to be 1,000-fold more resistant to inactivating effects of antibiotics (rifampicin, ciprofloxacin) when being protected inside macrophage and amoeba (Barker *et al*. 1995). In order to achieve killing of *L. pneumophila* antibiotics must be capable of penetrating into macrophage cells (Barker *et al*. 1995). The current antibiotics of choice for treatment of Legionnaires' disease are azithromycin, erythromycin and fluoroquinolones, although rifampicin is often also recommended (Fields *et al*. 2002). Aminoglycosides (such as gentamicin, kanamycin and streptomycin), chloramphenicol and cefoxitin are also active against *L. pneumophila* at low concentrations (Thornsberry *et al*. 1978; Edelstein & Meyer 1980).

These antibiotics target cytoplasmic components of the bacteria. Azithromycin and erythromycin bind to the 50S subunit of the bacterial ribosome and inhibits translation of mRNA (Zuckerman 2004). Fluoroquinolones inhibit DNA replication by targeting DNA gyrase or topoisomerase IV in the enzyme-DNA complex (Hooper 2001). Aminoglycosides inhibit protein synthesis (Kotra *et al*. 2000). Rifampicin binds to DNA-dependent RNA polymerase and inhibits bacterial DNA-dependent RNA synthesis (Hartmann *et al*. 1985).

However, the emergence of drug-resistance of *L. pneumophila* has been reported (Moffie & Mouton 1988; Fong *et al*. 2010; Ferhat *et al*. 2009), urging the need for discovery of novel drug classes.

1.3 Pathogenic pathway of *Legionella* bacteria

Legionella bacteria are ubiquitous in freshwater environments (Fliermans *et al*. 1981). *L. pneumophila* has been detected in a wide range of fresh water sources, from natural aquatic to human-made water systems (Fields *et al*. 2002). However, *Legionella* bacteria are not free-living aquatic bacteria. During their life cycle, they parasitize or form a commensal relationship with amoebae (Figure 1.1) (Rowbotham 1980; Fields 1996; Kwaik et al. 1998; Newton et al. 2010).

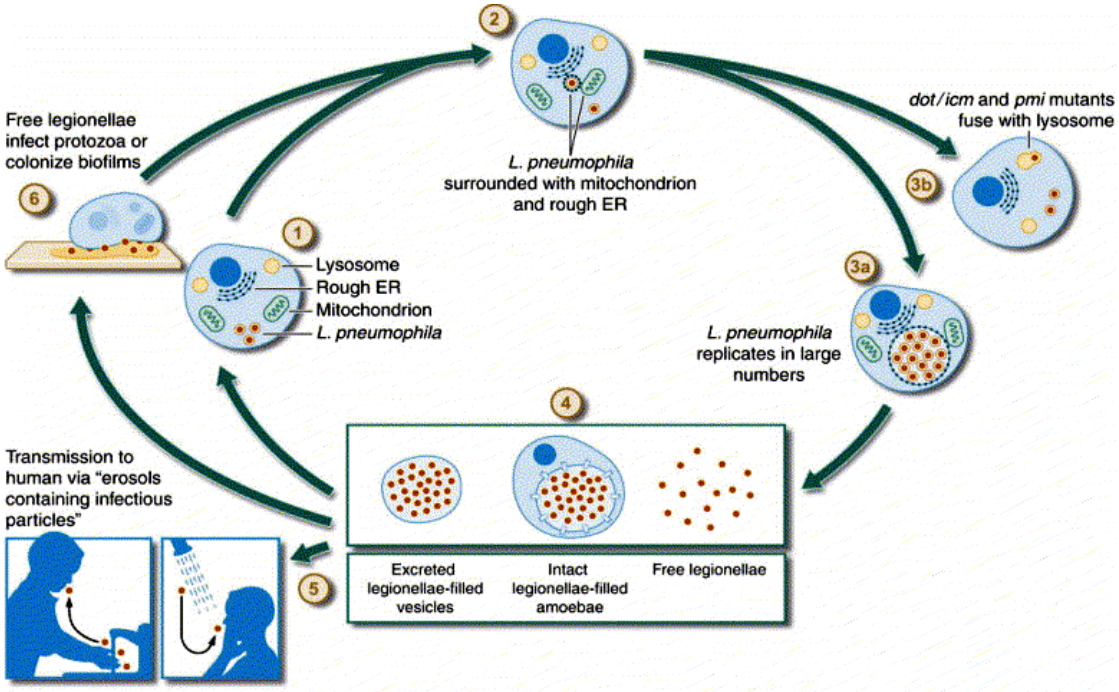


Figure 1.1: Life cycle of *L. pneumophila* within protozoa. (1) *L. pneumophila* infects protozoa. (2) Upon infection, *L. pneumophila* locates in a membrane-bound vacuole and is surrounded with host cell organelles, such as mitochondria and rough ER. *L. pneumophila*-containing vacuole does not fuse with lysosomes. (3) 3a. *L. pneumophila* replicates in large numbers. 3b. Some mutants fuse to lysosomes and undergo lysosomal degradation. (4) Different forms of the infectious particle: excreted *legionella*-filled vesicles, intact *legionella*-filled amoebae, or free legionella that have lysed their host cell. (5) Transmission to human via inhalation of aerosols containing infectious particles. (6) Leaving the host cells, free *legionella* reinfect other protozoa, or recolonize biofilms. Reprinted from (ASM News. 66(10) (2000): 609-616) (Newton *et al*. 2010).

Legionella infection usually occurs through inhalation of contaminated aerosols produced by human-made water systems and faucets (Albert-Weissenberger *et al*. 2007). Cooling waters are the frequent sources in reported community-acquired outbreaks (Minh *et al*. 2006). Potable water may also be an important source of sporadic and epidemic legionellosis (Fraser 1985). Other modes of transmission of *Legionella* are aspiration and direct instillation into the lung (RR *et al*. 1986). However, person-to-person transmission has never been reported (Newton *et al*. 2010).

The interaction of *L. pneumophila* with eukaryotic cells is similar to its interaction with environmental hosts. After inhalation of contaminated aerosol droplets, the bacteria reach the alveolar parts of the lungs where they are phagocytosed by alveolar macrophages. Uptake of *L. pneumophila* by

monocytes and macrophages has been shown to occur through coiling phagocytosis (Horwitz 1984) or conventional phagocytosis (Molmeret *et al*. 2004). The microbes can block the fusion of lysosomes with the phagosome, preventing the acidification of its phagosome (Horwitz & Maxfield 1984; Horwitz *et al*. 1983) and therefore escape the phagosome-lysosome degradation pathway. Once inside the macrophages or monocytes, the *L. pneumophila* phagosome is surrounded by host cell organelles such as mitochondria and vesicles (Horwitz *et al*. 1983). The bacterium begins to replicate until the nutrients of the host macrophage are consumed. The newly born bacteria move into their transmission phase, lyse the host macrophage (Molmeret *et al*. 2010) and restart the infection cycle. The process of lysing the host cells leads to the destruction of lung cells (Figure 1.2), resulting in Legionnaires' disease (Winn & Myerowitz 1981).

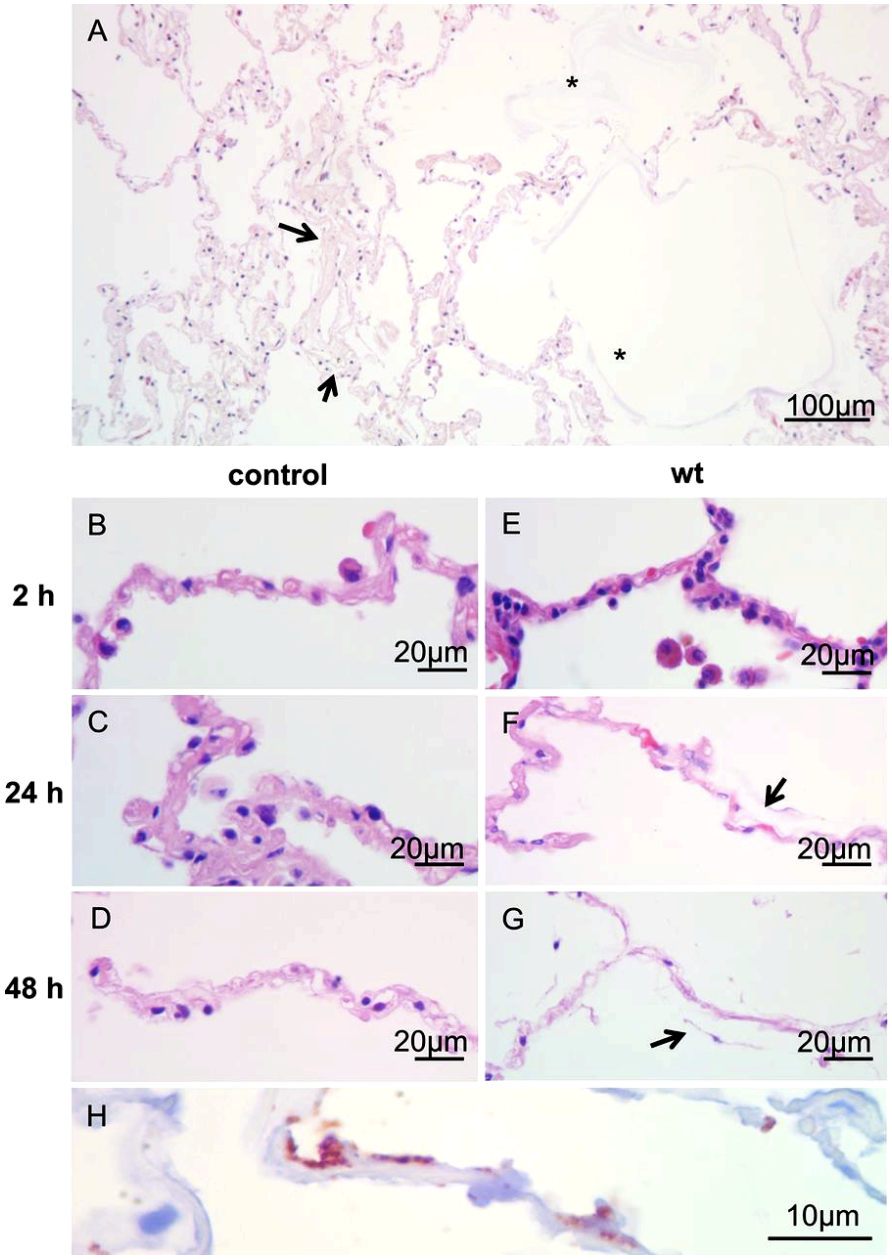


Figure 1.2: Tissue destruction caused by *L. pneumophila* within the human lung over time, visualized by Hematoxylin-eosin stain. A: A human lung tissue explant (HTLE) infected by *L. pneumophila* after 48 h. It was observed that collagen fibers within the alveolar septa appeared loose (arrow), indicating decreasing septal integrity. B, C, and D: Uninfected control HTLEs after 2, 24, and 48 h of incubation, respectively. E, F, and G: HTLEs infected by *L. pneumophila* at 2, 24, and 48 h after infection, respectively. Integrity lost of connective tissue, delamination of alveolar epithelial cells and protein exudate in the alveolar lumen (*) are phenomena observed in infected tissue, in comparison with uninfected controls. After 48 h of incubation, *L. pneumophila*-infected tissue exhibits severe damage (G). H: Using anti-Mip antibody to visualize *L. pneumophila*, severe tissue damage at the sites where the bacteria located (Jäger *et al*. 2014).

Various virulence factors and machineries of *L. pneumophila* participating in cytopathology of Legionnaires' disease have been described, including protein secretion systems as well as cytotoxic or hemolytic factors. Protein secretion systems are in charge of delivering virulence proteins of the bacteria passing through bacterial inner and outer membranes to their specific place of action (De Buck *et al*. 2007). For example, the Icm/Dot system involves in the promotion of phagocytosis, creating a nutrient-rich organelle that escapes the lysosomal degradation pathway, preserving the integrity of the *L. pneumophila* phagosome, the induction of apoptosis and lysis of the host cells (Molmeret *et al*. 2007). Several virulent proteins involved in pathogenesis of *L. pneumophila* have been identified including proteases, pore-forming toxins and bacterial lipolytic enzymes. Among these factors, bacterial phospholipases play important roles in the bacterial pathogenesis. Their enzyme activities result in destruction of the host cell membrane, or in products that manipulate host cell functions in a way that benefits the bacteria. Details on their functions and role in bacterial pathogenesis will be discussed in the next section.

1.4 Functions and role of the bacterial phospholipases in virulence

Phospholipases hydrolyze ester bonds of phospholipids (PL) into fatty acids and other lipophilic substances. There are 4 groups, termed phospholipases A-D, attacking different positions of the phospholipid (Figure 1.3). Phospholipase A (PLA) hydrolyses the carboxylester bonds at the sn-1 or sn-2 position, releasing fatty acids together with a lysophospholipid. This lysophospholipid can be further cleaved by a lysophospholipase A (LPLA), resulting in glycerophospholipid (Figure 1.4). Phospholipase B (PLB) acts on both fatty acid residues. Phospholipase C (PLC) and phospholipase D (PLD) attack the glycerol oriented and the alcohol-oriented phosphodiester bond, respectively. By cleaving components of lipid bilayers and releasing bioactive molecules, the bacterial phospholipases are involved in many disease-promoting processes from membrane destruction to interference with host cell signaling pathways (Figure 1.3) (Schmiel & Miller 1999; Istivan & Coloe 2006; Lang & Flieger 2011; Kuhle & Flieger 2013).

In recent years, an increasing number of virulence factors with PLA activity involved in events ranging from invasion into host cells to severe lung disease have been discovered (Banerji *et al*. 2008). PLA activity has been shown to

play the major role in lung surfactant destruction and hemolysis by *L. pneumophila* (Flieger, Gong, *et al*. 2000; Flieger, Gongab, *et al*. 2000). Up to date, 15 potential and confirmed PLA or LPLA enzymes have been identified. These proteins are either secreted into the culture medium or associated with the bacterium (Lang & Flieger 2011).

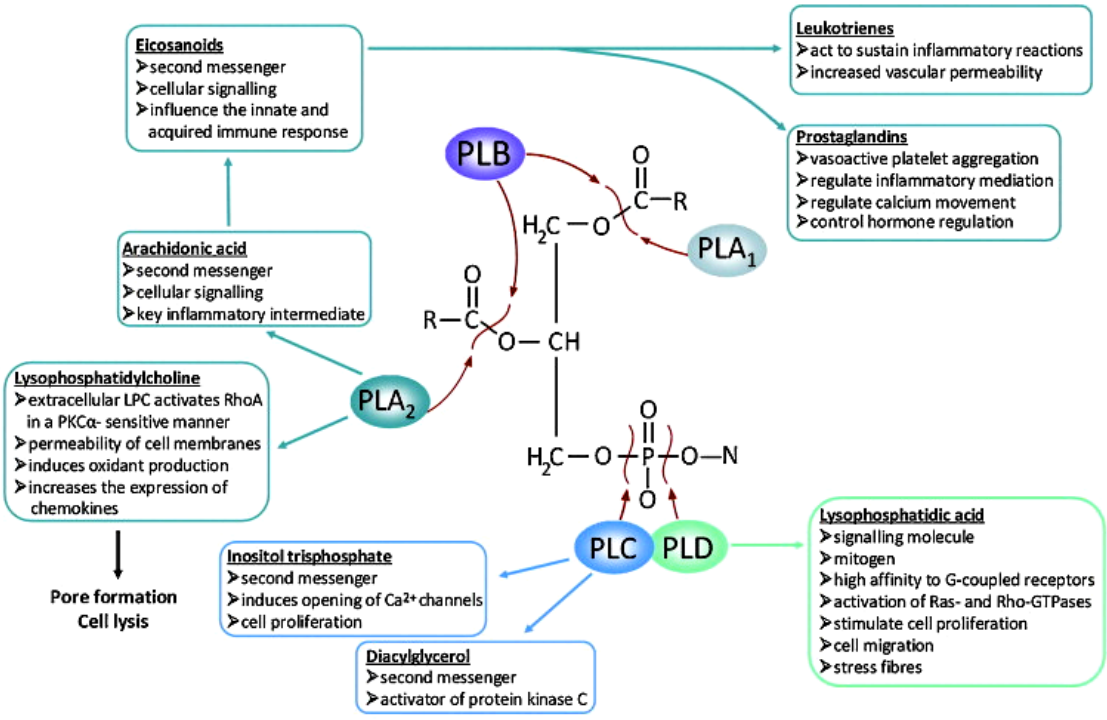


Figure 1.3: Possible impacts of phospholipases on host cells. For example, PLA_2 cleavages of phostidylcholine results in lysophosphatidylcholine which possesses a variety of functions within eukaryotic cells that eventually leads to pore formation and cell lysis. Archidonic acid or 1,2-diacylglycerol derived from a phospholipid as a result of PLA_2 or PLC activity, respectively, can act as second messengers involving in cellular signalling events. PLD activity results in lysophosphatidic acid that acts as as a lipid mediator. In particular, it plays as an inducer of cell proliferation, migration and survival. It also activates various signal transduction pathways, including those initiated by Ras- and Rho- GTPases, by binding to G-coupled receptors with high affinity (Lang & Flieger 2011).

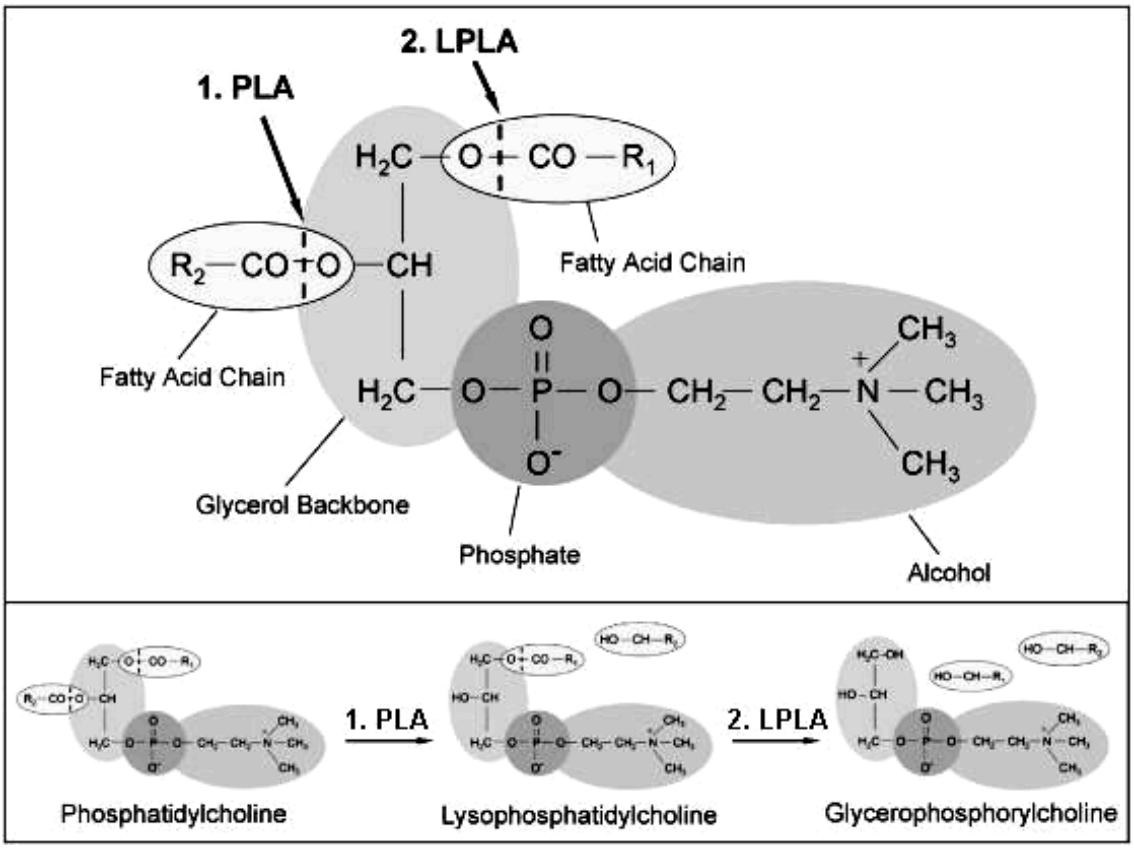


Figure 1.4: Two-step hydrolysis by *L. pneumophila* phospholipase A (PLA) and lysophospholipase A (LPLA). In the first hydrolysis step, *L. pneumophila* exhibits PLA activity which cleaves one fatty acid residue from a phospholipid (phosphatidylcholine) resulting in the cytotoxic agent lysophospholipid (lysophosphatidylcholine). In the second hydrolysis step, LPLA cleaves the second fatty acid chain from lysophospholipid and releases glycerophospholipid (glycerophosphatidylcholine) (modified from Banerji *et al*. 2008).

1.5 PlaB – an important virulence factor of *L. pneumophila*

Among phospholipases A and lysophospholipases A of *L. pneumophila*, PlaB is the first enzyme confirmed to be pathogenic, involved in replication in lung cells and dissemination of the bacteria in an *in vivo* guinea pig model (Schunder *et al*. 2010). In this model, the guinea pigs were infected with either wild-type or a *plaB* mutant *L. pneumophila* strain. Two days after infection, the number of bacteria growing in lung and spleen was counted. It was observed that while the wild-type bacteria increased 400-fold in the lung, the *plaB* mutant bacteria increased only 20 fold. The dissemination of the plaB mutant bacteria to the spleen was 100 fold less than the wild type. Histopathology data of guinea pig lung tissue showed much more prominent inflammation, infection and tissue destruction in lungs infected by the wild type strain than in ones infected by the

plaB mutant strain. The endothelium of the lungs infected with the *plaB* mutant was nearly intact while accumulated endothelial debris blocking the alveoli was observed in lungs infected with the wild-type bacteria (Schunder *et al*. 2010).

PlaB has both phospholipase and lysophospholipase activities and additional hemolytic potential. The enzyme mainly hydrolyses long-chain fatty acid substrates (PC and PG) with preferences for lipids containing fatty acyl residues larger than eight carbon atoms, as well as the resulting lysophospholipids (Flieger *et al*. 2004; Bender *et al*. 2009). DPPC, a main component of eukaryotic membrane and lung, is the preferred substrate of *L. pneumophila* PlaB (Flieger *et al*. 2004). This perfectly fits with the observation that *Legionella* occupy lung cells and lead to cell membrane destruction. Interestingly, the specificity for PC hydrolysis is vital for PlaB hemolytic activity. Mutations that resulted in loss of ~90% efficiency of PC cleavage also led to loss of hemolysis (Bender *et al*. 2009).

PlaB is the most prominent cell-associated phospholipase A of *L. pneumophila* and its active part is presented to the external environment. These conclusions were drawn from the following facts: the supernatant of *L. pneumophila* exhibited no PlaB-related PLA activity and a *plaB* mutant almost completely lost its cell-associated lipolytic activity; the hemolytic activity of PlaB was shown to be contact-dependent (Flieger *et al*. 2004); sucrose density gradient centrifugation revealed PlaB was located in the outer membrane fractions, and its active part is susceptible for proteinase K digestion, suggesting its active part is exposed on the surface of the bacteria (Schunder *et al*. 2010).

1.6 PlaB represents a novel lipase family

Functionally speaking, it is obvious that PlaB is a member of lipase families. With around one hundred three-dimensional lipase structures representing around 30 organisms, current knowledge on lipase structures indicates that the enzymes share a common α/β hydrolase fold and a conserved catalytic triad serine (S), aspartic acid (D) and histidine (H). The α/β hydrolase fold generally consists of eight β strands connected by six α helices (Figure 1.5). The catalytic triad mostly arranged within conserved amino acid blocks. For example, catalytic serine is embedded in the typical GXSXG in true lipases (Arpigny & Jaeger 1999) or in GDSL motif in lipase family II (Godoy *et al*. 2012). See table

1.1 for more details on conserved motifs in which each catalytic residue is embedded.

Figure 1.5: An example of the alpha/beta-hydrolase fold. Crystal Structure of a Thermostable Lipase from ***Bacillus stearothermophilus*** P1 (PDB id: 1ex9), α-helices are represented by red spirals, β-strands are indicated by arrows. Red dotted line represents for bacterial outer membrane (Tyndall et al. 2002).

Table 1.1: Conserved motif around catalytic triad residues of lipase families

Lipase family/Subfamily	Consensus motif of catalytic S	Consensus motif of catalytic D	Consensus motif of catalytic H
I	GXSXG	N(X)DGL(X)V	N(X)HL(X)D
II (GDSL)	GDSL	LFXDX	XHPT
III	N/A	N/A	N/A
IV (HSL)	GDSAGG	XDPL	XHGF
V	GXSMGG	GDXDX	DXGHX
VI	GFSQG	HGXXDXV	MGHE
VII	N/A	N/A	N/A
VIII	N/A	N/A	N/A
PlaB	THSTG	GSDGVV	SHS

Red: catalytic residue. N/A: consensus motif is not available (Adapted from alignment results of (Arpigny & Jaeger 1999))

In structural view, however, PlaB and its homologs do not belong to any known lipase families. Although the enzyme also possesses the “classical” catalytic triad of nucleophilic serine 85, aspartate 203 and histidine 251, protein blast suggested it does not exhibit significant sequence identity to known lipases (Flieger *et al*. 2004). Moreover, the protein presents some unique features (Bender *et al*. 2009). The first glycine of the consensus motif GXSXG in true lipases is replaced with a threonine residue in PlaB (Figure 1.4, Table 1.1). Replacing this residue with glycine reduces activity of PlaB by 95%, while a substitution with valine reduced the DPPC/MPLPC hydrolysis activity up to 50% (Bender *et al*. 2009). These results indicated that this residue is crucial for the enzyme activity. The two other crucial catalytic residues are embedded in unique motifs that were not found in other characterised lipase families (Bender *et al*. 2009). Following the catalytic histidine of PlaB is the polar amino acid serine rather than non-polar, hydrophobic residues that are usually present in other lipases (Flieger *et al*. 2004) (Figure 1.4, Table 1.1). It is suggested from previous studies that this serine residue may play a vital role in the expression of phospholipase A activity. Replacing serine in the immediate vicinity of catalytic histidine by valine in *Staphylococcus hyicus* reduced phospholipase activity more than 10-fold (van Kampen *et al*. 1998), while replacing the leucine adjacent to the active-site histidine with serine converted the lipase of *Bacillus thermocatenulatus* into a phospholipase A (Kauffmann & Schmidt-Dannert 2001). How exactly these differences in PlaB define its phospholipase activities remains unknown.

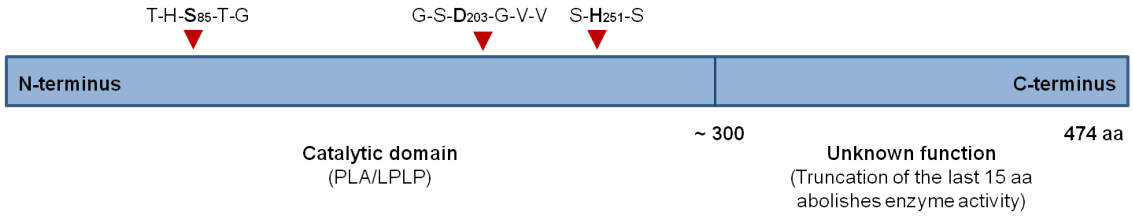


Figure 1.6: Structural scheme of PlaB. The three brown arrows indicate positions of catalytic residues (in bold) and the unique motifs in which they are embedded, respectively (Adapted from (Kuhle in RKI, personal communication).

Adjacent to the catalytic active domain of PlaB is a C-terminal extension of about 200 amino acids. The function of this domain is still unknown. Truncation of only 15 amino acids at the C-terminus is sufficient to stop the cell-associated phospholipase A activity, suggesting that the C terminus might be important for activation of PlaB (Flieger *et al*. 2004).

1.7 Structure-based drug design (SBDD)

Detail information on the active site and other important regions of an enzyme at atomic level enables a chemist to design small molecules that bind to one of these sites and interfere with the enzyme's activity (Beddell *et al.* 1976; Walsh 1983; McPherson 1999). These small molecules are modeled by computer modeling programs and quite often, they are analogs of natural substrates of the enzyme. Experimental determination of the enzyme structure is perhaps the most accurate way of obtaining the necessary information for inhibitor design, which includes the rigidity, geometry of the binding site and its hydrophobic character, electrostatic landmarks, the chemical features, and mechanisms related to biological functions of the enzyme (McPherson 1999; Halgren 2009; Villoutreix *et al*. 2007; Cheng *et al*. 2007).

Since the first published work describing SBDD for human hemoglobin in 1976 (Beddell *et al*. 1976), SBDD has become a key step in the discovery of many new and important marketed drugs. For example, structural determination of the HIV protease allowed the successful design of five protease inhibitors that are now commercially available (Erickson *et al*. 1990; Roberts *et al*. 1990; Istry *et al*. 1994).

In the field of antibacterial-drug discovery, various antibacterial drugs were developed and have had significant contributions for human society (Table I.2). However, the widespread global occurrence of bacteria resistant to the antibiotics and synthetic drugs has been emerging (Chopra 2013; Simmons *et al*. 2010; Agarwal & Fishwick 2010). Much effort has been spent to develop new drugs but they do not adequately meet growing medical needs (Chopra 2013; Agarwal & Fishwick 2010). Developing derivatives from older classes has a disadvantage that they are prone to existing bacterial resistance mechanisms. Thus, new drug classes targeting new virulence factors are urgently needed, and SBDD offers a promising approach that allows to discover new drugs rationally.

Table 1.2: Targets and classes of antibacterial drugs (Simmons *et al*. 2010)

Drug target	Drug classes
Cell wall synthesis	B-Lactams, bacitracin, cycloserine, fosfomycin and glycopeptides
Cell membrane integrity	Daptomycin and polymyxins
Nucleotide biosynthesis	Sulfonamides and trimethoprim
DNA replication	Quinolines, nitrofurans and nitromidazoles
RNA synthesis	Rifamycins
Protein synthesis	Aminoglycosides, chloramphenicol, fusidic acid, ketolides, macrolides, oxazolidinones, streptogramins, tetracyclines and mupirocin

1.8 PlaB is a promising drug target

Regarding drug delivery and targeting, PlaB represents a promising drug target because the enzyme is located outside the membrane and is hence more accessible as a drug target than other cytoplasmic proteins. Additionally, PlaB is crucial for bacteria growth in lung and dissemination of the bacteria to the spleens. Blocking PlaB can at least inhibit the dissemination of PlaB to other cells inside the host. It is hypothesized that PlaB may play some role in bacterial replication, too. If a drug can be designed in a way so it can penetrate macrophage, the inhibitor can then block the PlaB and therefore inhibit the bacterial replication inside lung cells. Furthermore, because PlaB is a novel lipase, the risk that the drug targets a human homolog is significantly reduced.

2 Aims of the study

Despite extensively biochemical studies in Antje Flieger's group at Robert Koch Institute (Wernigerode) where the enzyme was first identified and characterised, there remain many questions regarding the enzyme functions. For examples:

1. How does structure of PlaB look like?
2. How is PlaB activated and regulated?
3. Is there any relevance between the unique amino acid environment around the catalytic triad of PlaB with its function and its ability of exhibiting both PLA and LPLA activity?
4. Is metal required for enzyme activities?
5. What is the function of the C-terminal domain?
6. Why does PlaB favor long-chain phospholipids?
7. Can the structure of PlaB be used as a template to design a PlaB inhibitor?

The aim of this thesis was thus the crystallization and X-ray structure determination of PlaB in order to address the above-mentioned questions.

A structure of PlaB will give us a detailed picture of the overall PlaB architecture, and will provide structural insights into its enzymatic activity. A structure of the enzyme in complex with its substrates will give invaluable information to chemists in finding new inhibitors. With atomic details of the substrate bound to the protein in the catalytic pocket, chemists may predict and synthesize homologs of substrates that can block the activity of the protein. Thus, the structures of both apo and substrate-bound form of PlaB can give us more knowledge about the function of C-terminal part of the protein.

3 Materials and Methods

3.1 Materials

The chemicals were purchased with the quality standard “pro analysis” (p.a.) from the following companies: Eurofins MWG Operon, Fermentas, Fluka, GE-Healthcare, Hampton Research, Invitrogen, Merck, Millipore, Qiagen, Roche, Roth, Sigma-Aldrich and Stratagene.

3.1.1 Enzymes, standards

Table 3.1: Enzymes and standards used in this work

Enzyme	Company
DNase I	Roche
Standard	
Smart Ladder (DNA)	Eurogentec
Precision Plus Protein All Blue Standards (Protein)	BIO-RAD

3.1.2 Kits

Table 3.2: Commercial kits used in this work

Kit	Company
QIAprep® Spin Miniprep Kit	Qiagen
QuikChange® Site-Directed Mutagenesis Kit	Stratagene
Proteolysis screen	Hampton Research

3.1.3 Commercial crystallization screens

The following crystallization screens from Qiagen were used for screening initial crystallization conditions:

Table 3.3: Commercial crystallization screens used for finding initial crystallization conditions

The $AmSO_4$	JCSG+
The Anions	JCSG Core I
The Cations	JCSG Core II
The Cryo	JCSG Core III
The Mb Class	JCSG Core IV
The MPD	The Classics
The PEGs I	The pH Clear
The PEGs II	

The Morpheus screen from Molecular Dimensions, as well as the Additive Screen and Silver bullets from Hampton Research were also used.

3.1.4 Plasmids

Table 3.4: Plasmids used in this work

Plasmid	Derivation	Reference
pGP172-PlaB	PlaB in pGP172	Kuhle *et al.* 2014
pGP172-PlaB D203N	Site directed mutagenesis of PlaB in pGP172	Kuhle *et al.* 2014
pGP172-PlaB D203N_459	Truncation of the last 15 aa of PlaB D203N mutant in pGP172	Kuhle *et al.* 2014

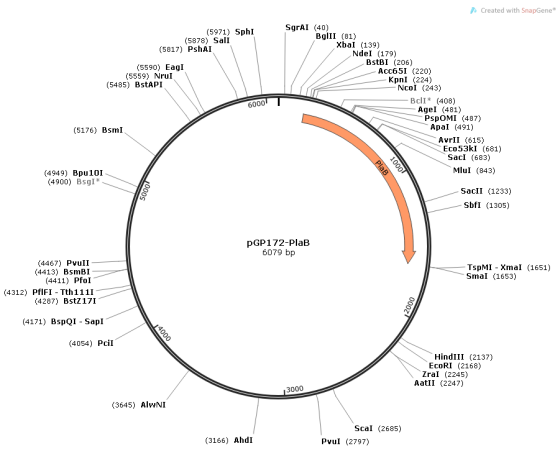


Figure 3.1: Map of pGP172-PlaB plasmid. The plasmid carries a T7 promoter which is activated by IPTG, and an ampicillin resitance gene for selection (Kuhle *et al.* 2014).

3.1.5 Bacterial strains

Table 3.5: Bacterial strains used in this work

E.coli strains	Genotype	Source
Top10	F ¯, mcr A Δ(mrr -hsd RMS-mcr BC) Φ 80 lac Z Δ M15 Δlac X74 recA1 araD139 Δ(ara-leu)7697 gal U gal K *rps*L *end*A1 *nup*G	Invitrogen
XL1-Blue supercompetent cells	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´proAB lacIqZΔM15 Tn10 (Tetr)]*	Stratagene
Tuner (DE3)	*F ¯lacY ompT gal dcm lon hsdSB(r $_B$¯ m $_B$¯) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*	Novagen

The first two strains were used for DNA preparation while the last strain was used for protein expression.

3.1.6 Culture media and supplements

Table 3.6: Compositions of culture media and supplements

Type of medium	Composition (in 1 L media)
Luria Bertani (LB)	10 g tryptone, 5 g NaCl, 5 g yeast extract
Terrific Broth (TB)	12 g tryptone, 24 g yeast extract
SOC-Medium	20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 2.5 mM KCl, 10 mM MgCl2, 20 mM glucose
Minimal medium*	20.8 mM NH_4Cl, 24.5 mM KH_2PO_4, 24.9 mM Na_2HPO_4
SeMet-nutrients*	1.11 M D-Glucose, 24.7 mM $MgSO_4$, 332 µM Thiamine HCl, 260 µM $Fe_2(SO_4)_3$
Amino acid mix	100 mg Lysine, 50 mg Isoleucine, 50 mg Leucine, 50 mg Valine

*Although the final concentration of each composition is calculated for 1 L media, the minimal medium was prepared with volume of 900 ml, and SeMet nutrients were prepared in 100 ml. All of the media were autoclaved before using. SeMet nutrients and amino acid mix were prepared fresh.

3.1.7 Buffers and reagents

Table 3.7: Buffer compositions

Type of buffer	Composition
DNA loading buffer (10x)	10 mM Tris-HCl, pH 7.5, 0.05% (w/v) bromphenol blue, 1 mM EDTA, 50% glycerol
TAE buffer	40 mM Tris-HCl pH 8.2, 20 mM CH_3COONa, 1 mM EDTA
Lysis and wash buffer	100 mM Tris pH 8.0, 100 mM NaCl, 1mM EDTA, 5 mM DTT
Elution buffer	100 mM Tris pH 8.0, 100 mM NaCl, 1 mM EDTA, 5 mM DTT, 8 mM d-Desthiobiotin
Regeneration buffer	100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 1 mM HABA
SEC buffer	20mM Tris-HCl pH 8, 100mM NaCl, 5mM DTT
Ion exchange buffer A	20 mM HEPES pH 7.0, 100 mM NaCl, 5 mM DTT
Ion exchange buffer B	20 mM HEPES pH 7.0, 1 M NaCl, 5 mM DTT
SDS-PAGE lower buffer (4x)	1.5 M Tris-HCl, pH 8.8
SDS-PAGE upper buffer (4x)	0.5 M Tris-HCl pH 6.8, 0.4% (w/v) SDS
SDS-PAGE running buffer	25 mM Tris-HCl, 192 mM glycine, 0.1% (w/v) SDS
SDS-PAGE sample buffer (2x)	1 ml 1 M Tris-HCl pH 6.8, 2.4 mM glycerol, 0.8 g SDS, 2 mg Coomassie blue G-250, 0.31 g DTT, add H_2O to 10 ml
SDS-PAGE separating gel	4.7 ml H_2O, 10 ml Polyacrylamide solution (33.3%), 5 ml 4x Lower buffer, 0.2 ml SDS (10%), 20 µl TEMED, 50 µl APS (25%)
SDS-PAGE stacking gel	5.9 ml milliQ water, 1.5 ml Polyacrylamide solution (33.3%), 2.5 ml 4x Upper buffer, 15 µl TEMED, 25 µl APS (25%)
Acidic native stacking gel (pH 6.8)	3% Acrylamide + Bis-Acrylamide, 62.5

	mM KOH – CH_3COOH, milli-Q water, 0.05% APS, 5 μl TEMED (for 5 ml gel)
Acidic native seperating gel (pH 4.5)	12% Acrylamide + Bis-Acrylamide, 375 mM KOH - CH_3COOH, milli-Q water, 0.05% APS, 5 μl TEMED (for 5 ml gel)
Neutral native stacking gel (pH 8.0)	3% Acrylamide + Bis-Acrylamide, 62.5 mM KOH-MOPS, milli-Q water, 0.05% APS, 10 μl TEMED (for 5 ml gel)
Neutral native seperating gel (pH 6.8)	12% Acrylamide + Bis-Acrylamide, 375 mM KOH-MOPS, milli-Q water, 0.05% APS, 10 μl TEMED (for 5 ml gel)
1x electrode buffer for acidic native gel	350 mM β-alanine, 140 mM CH_3COOH, milli-Q water
1x electrode buffer for neutral native gel	100 mM Histidine, 22 mM MOPS, milii-Q water
1x sample buffer for acidic native gel	30% Glycerol, 188 mM KOH-CH_3COOH, milli-Q water, 0.1% Bromophenol blue
1x sample buffer for neutral native gel	30% Glycerol, 188 mM KOH-MOPS, milli-Q water, 0.1% Bromophenol blue

3.2 Methods

3.2.1 Standard methods

Methods used in this study are adapted from collections of methods and protocols (Sambrook, J. & Russell, D.W., 2001). Only variations of standard protocols are described below.

3.2.2 Protein production and purification

3.2.2.1 Expression of native protein

The following procedure is for 1 L of cell culture. Typically 50 µg DNA plasmid was transformed into competent Tuner (DE3) cells. Single colonies were selected on an Ampicillin-containing LB-plate. One fresh colony was then used for inoculation of an overnight culture in 50 ml LB containing 100 µg/ml Amp at 37°C, with shaking at 160 rpm. 10 ml overnight culture was used to inoculate 1 L TB media containing antibiotics. The culture was grown at 37°C at 140 rpm until OD_{600} reached 0.6. Protein expression was induced with the addition of 0.5 mM IPTG and the culture was incubated overnight at 20°C. Cells were harvested by centrifugation at 6,000 rpm at 4°C for 15 minutes and immediately lysed for protein purification.

3.2.2.2 Expression of SeMet-labeled protein

For the production of Se-Met labeled protein, the final protocol is as following: A fresh colony was used for inoculation of an overnight culture in 30 ml LB containing antibiotics at 37°C. 10 ml overnight culture was used to inoculate 1 L LB media containing antibiotics. When the OD_{600} reached 3.3, the culture was divided into 4 parts. The cells of each part were harvested by centrifugation at 6,000 rpm for 15 minutes, the supernatant was removed, and the cell pellet was washed in 200 ml milli-Q water. The cells were then re-suspended in minimal medium supplemented with media nutrients, using a volume to yield an OD_{600} of 0.6. Amino acid mix, 60 µg/ml SeMet and 0.5 mM IPTG were added subsequently. The cells were incubated overnight at 20°C and afterwards harvested by centrifugation.

3.2.2.3 Protein purification

After resuspension in 30 ml lysis buffer containing DNase I, 20 mg lysozyme and 1 tablet of Complete Protease Inhibitor Cocktail, the cells were lysed by passing them two times through a homogenizer at 20 kpsi at 4°C. The lysate was cleared by centrifugation at 16,000 rpm for 45 min at 4°C. The supernatant was loaded onto a Strep-Tactin column pre-equilibrated with Wash buffer. Unspecifically bound protein was removed by washing the column with 5 column volumes of Wash buffer. The target protein was eluted with 6 x 0.5 CV Elution buffer. The protein content of the eluted fractions was measured at 280 nm and the purity was determined by 12% SDS-PAGE. Fractions with high enough protein content and purity were applied to reverse anion exchange. Fractions from the flow-through were collected, analyzed by SDS-PAGE and the fractions containing the pure protein were pooled and concentrated to 5 mg/ml. Vivaspin concentrators with a molecular weight cutoff (MWCO) of 10 kDa were used for concentration step.

3.2.3 Protein analytical methods

3.2.3.1 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was used to monitor the purity and yield during the protein purification (Laemmli 1970). The gels were composed of two parts: the upper stacking gel (5% acrylamide) and the lower resolving gel (12% acrylamide). 15 µl of protein solution together with 5 µl of SDS loading buffer (containing DTT) were heated for 5 min at 95°C and then run at constant 40 mA per gel on an SDS gel until the dye front reached the end of the gel (usually 40 min). After the electrophoresis, the gel was washed with water to remove SDS and then stained with Instant Blue Coomassie dye for 10 min.

3.2.3.2 Mass spectrometry (MS)

The identity and the correct labeling of the produced proteins were confirmed by MS using either MALDI-TOF or ESI-MS. These analyses were carried out by Dr. Manfred Nimtz and Undine Felgenträger (HZI, Braunschweig).

3.2.3.3 Dynamic light scattering

All purified protein samples were analyzed by DLS to assess their polydispersity (Proteau et al. 2010; Borgstahl 2007). The method was also used to examine the change of the oligomerization state of the protein at different concentrations. To remove large particles disturbing measurements, the protein solutions were centrifuge at 14,000 rpm for 20 min and 15 µl sample volumes were transferred in a dust-free DLS-cuvette. The experiments were carried out at 4°C and room temperature on a DynaPro Titan Temperature Controlled MicroSampler instrument (Wyatt Technologies). For each sample, 25 individual experiments were recorded.

3.2.3.4 Limited proteolysis screen

Proti-Ace Kit (Hampton Research) was used to determine stable sub-domains suitable for crystallization. The kit consists of these below proteases:

1. alpha Chymotrypsin (α-C): a serine endopeptidase that selectively cleaves peptide bonds formed by aromatic residues (tyrosine, phenylalanine and tryptophan).
2. Actinase E (A-E): a serine protease displaying a wide range of substrate speficities. However, it prefers to hydrolyze peptide bonds on the carboxyl side of glutamic or aspartic acid.
3. Bromelain (BR): a cysteine endopetidase that breaks peptide bonds of non-terminal amino acids.
4. Clostripain (CL): a protease that cleaves proteins on the carboxyl peptide bond of arginine.
5. Endoproteinase Glu-C (EG-C): a serine protease which hydrolyzes peptide bonds at the carboxyl side of glutamyl and aspartyl residues. The specificity of EG-C is dependent upon the buffer and pH employed as well as the structure around the potential cleavage site. No cleavage will occur if a proline residue is on carboxyl side. The enzyme also exhibits esterase activity.
6. Elastase (EL) cleaves peptide chains mainly at the carboxy side of small, hydrophobic amino acids such as glycine, alanine and valine.
7. Papain (PA) exhibits broad specificity, cleaving peptide bonds of basic amino acid, leucine, or glycine. It also hydrolyzes esters and amides, exhibits a preference for an amino acid bearing a large hydrophobic side chain at the P2 position. It does not accept Val at the P1' position.

8. Pepsin (PE) is most efficient in cleaving peptide bonds between hydrophobic and preferably aromatic amino acids such as phenylalanine, tryptophan and tyrosine.
9. Proteinase K (P-K) is a serine protease. The predominant site of cleavage is the peptide bond adjacent to the carboxyl group of aliphatic and aromatic amino acids with blocked alpha amino groups.
10. Subtilisin (SU) is not specific in activity.
11. Thermolysin (TH) specifically catalyzes the hydrolysis of peptide bonds containing hydrophobic amino acids.
12. Trypsin (TR) cleaves peptide chains mainly at the carboxyl side of the amino acids lysine or arginine, except when either is followed by proline.

After 4 hours and after overnight incubation, aliquots were taken and proteolysis was stopped by mixing with SDS-PAGE sample buffer and heating for 60 sec at 95°C. The products of the reactions were analyzed by SDS-PAGE.

3.2.3.5 Lysine methylation

The methylation reaction was performed overnight in 50 mM HEPES (pH 7.5), 250 mM NaCl at protein concentration of 1 mg/ml. Twenty micro liters freshly prepared 1 M dimethylamine-borane complex (ABC) and 40 µl 1 M formaldehyde were added per ml protein solution, and the reactions were gently mixed and incubated at 4°C for 2 hr. A further 20 µl ABC and 40 µl formaldehyde were added and the incubation continued for 2 h. Following a final addition of 10 µl ABC, the reaction was incubated overnight at 4°C. Precipitated protein was removed by centrifugation before purification of the soluble methylated protein by size-exclusion chromatography in 50 mM Tris-HCl (pH 7.5), 200 mM NaCl. Appropriate peak fractions were pooled and concentrated, exchanged to 20 mM HEPES buffer (pH7.0), 100 mM NaCl in centrifugal concentrators to 5 mg/ml.

3.2.4 Protein crystallization

Concentrated and purified proteins were centrifuged at 14,000 rpm at 4°C for 20 min before crystallization. All proteins were crystallised by hanging- and sitting-drop vapor diffusion methods at constant temperatures (4°C, 12°C, 20°C and 30°C).

3.2.4.1 Initial screening

To screen for initial crystallization conditions, 0.2 µl of the concentrated, pure protein (typically 5 mg/ml) was mixed with the same volume of reservoir buffer of commercially available screens in a 96-well sitting drop vapor diffusion plate using the Honeybee 961 crystallization robot (Zinsser Analytics). The plate was sealed using a MancoTM Crystal Clear tape (Jena Bioscience). After one day and then every two days the plates were checked for the appearance of crystals with a microscope.

3.2.4.2 Optimization – external parameters

Initial crystals were evaluated and then further optimized in 24-well hanging drop vapor diffusion plates. External parameters that were used for optimization were protein and precipitant concentrations, pH, types of PEG, drop sizes and incubation temperatures. Crystal growth promoting additives were also tested using the Additive Screen and the Silver bullet screen (Hampton Research).

3.2.4.3 Optimization – chemical modifications

Chemical modifications were employed after exhausting above-mentioned approaches. Modification approaches included lysine methylation, limited proteolysis screening and C-terminal truncations to produce chemical modified proteins for initial crystallization screening.

3.2.4.4 Cryo-protection

During data collection, high-energy X-rays induce the formation of radicals within the crystal, resulting in crystal damage and thus rapid loss of diffraction (Garman 2010). This issue can be overcome by freezing the crystals in liquid nitrogen and collecting data in a stream of gaseous nitrogen held at 100°K. Yet, protein crystal and the surrounding mother liquor have high water content, freezing the crystal results in ice formation, which can damage the crystal and/or reduce the quality of the measured data by its own diffraction (ice rings). This is avoided by the addition of or co-crystallization with cryo-protectants. Several cryoprotectants were tested in this work, including glycerol, PEG 400, PEG 200, and paraffin. Glycerol was the best cryoprotectant. A solution of reservoir buffer including 20% glycerol was prepared, the crystal transferred

from its drop into this solution and then quickly frozen in liquid nitrogen.

3.2.5 Heavy atom screen by native gel

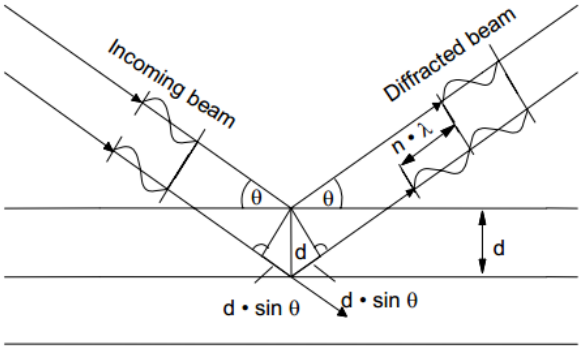
Native gels were prepared manually. The heavy atom at stock concentration of 100 mM was diluted to 10 mM, then 2 µl of heavy atom was taken and mixed with 2 µl of protein and incubate at room temperature for 1 h. After that, the mixture was applied to native gel for 4 h at 70 V.

3.2.6 X-ray analysis

3.2.6.1 X-ray measurement

X-ray crystallography is a fundamental technique to determine protein structures at atomic resolution. Structural details at atomic level allow elucidation of enzymatic mechanisms, molecular interactions and design of protein inhibitors. To determine protein structure by X-ray crystallography, a single crystal with sufficient size is required. The setup and mechanism of the X-ray diffraction analysis in this work are as follows: The crystal were frozen and irradiated with monochromatic X-rays, generated either by a rotating anode X-ray generator (home source) or a synchrotron. The crystals were kept at 100°K all time during the measurement by a continuous stream of cold nitrogen gas. Upon interaction with the protein crystal, X-rays are diffracted and the diffraction is measured on a CCD device. The crystal is rotated around one axis in small step (0.5°) and one image is recorded for each step. These images are used to calculate the intensity and position of the reflections, which are input parameters for the determination of crystal structure.

In which direction and under which angle an X-ray is diffracted is determined by Bragg`s Law (Figure 3-1). Knowing λ and θ, d is then computable. From d, lattice parameters can be calculated.



$$n \bullet \lambda = 2 \bullet d \bullet \sin \theta$$

Figure 3.2: Diffraction of X-rays from two parallel planes following Bragg's Law. n is an integral number, λ is the wavelength of the X-ray beam, d is the distance between two crystal planes and θ is the angle between incoming beam and crystal plane, usually also the crystal surface.

3.2.6.2 Structural analysis

The intensities and phase information of reflections are input parameters for the determination of crystals structure. Intensity of each diffraction spot should be strong enough for thoroughly analysis. However phase information is lost during measurements of protein crystals. Thus, methods to indirectly determine protein structures have been developed, including molecular replacement (MR) (Evans & McCoy 2007; Dodson 2007, Rossmann 2001; Rossmann & Blow 1962), single/multiple wavelength anomalous dispersion (Hendrickson & Ogata 1997), single/multiple isomorphous replacement (Green *et al.* 1954, Perutz 1956; Blow & Rossmann 1961), or in some cases combination of these methods are utilized.

3.2.6.3 Twin crystal analysis

In this work, twinning crystals were detected by L test (Yeates 1997) which considers pairs of locally related reflections instead of twin related reflections. The statistic L can be evaluated without prior knowledge of the twin law (Rupp 2009).

3.2.7 Surface site mutagenesis

The plasmid was used as the template for the PCR to generate double or triple mutants following the standard QuikChange protocol for site-directed mutagenesis (Strategene) with some modifications as below:

For mutant strand synthesis reaction, the sample mixture consisted of: 1x Pfu Ultra II buffer, 200 µM dNTP mix, 100 ng DNA templage, 0.5 µM forward primer, 0.5 µM reverse primer and 2.5 U Pfu Ultra II DNA polymerase.

Table 3.8: Cycling parameters for the site-directed mutagenesis method

Segment	Cycles	Temperature	Time
1	1	95℃	3 minutes
2	15	95℃	1 minutes
		52℃	1 minutes
		68℃	12.4 minutes
3	1	68℃	1 hour

After the amplification, 10 U DpnI was added into the mixture. The reaction mixture was then incubated immediately at 37℃ for 1 h to digest supercoiled dsDNA. The DpnI-treated DNA was transformed to XL1-blue competent cells. On the next day, 8 single clones were inoculated and subjected to DNA plasmid purification and sequencing to verify the correct mutations.

3.2.8 Bio-informatic tools

The Protparam tool (Gasteiger *et al*. 2003) was used to obtain theoretical data about the protein constructs such as extinction coefficients. Swiss-Model (Arnold *et al*. 2006), FUGUE (Shi *et al*. 2001), NCBI protein blast and Phyre (Kelley & Sternberg 2009) were used to search for structural homologs of PlaB. Sequence and secondary structure alignments were generated by NPS@ (Deléage *et al*. 1997).

4 Results

4.1 Development of PlaB purification scheme

4.1.1 Initial PlaB purification protocol

An initial purification procedure for 1-2 L bacterial cell culture expressing PlaB protein was established by Katja Kuhle as part of her Ph.D thesis. Following this protocol, the plasmid was transformed into BL21 cells and the protein was produced after induction with 2 mM IPTG overnight at 20°C. After cell lysis, the protein was purified from the supernatant via Strep-tag affinity chromatography and the purity was analysed by SDS-PAGE. The protein was almost over 90% pure but some impurities still remained (Figure 4.1A). The major band at 53 kDa was excised from the gel and confirmed as PlaB by mass spectrometry. Other bands were identified as *E.coli* proteins.

4.1.2 Conventional chromatographies

As impurities still remained visible on SDS-PAGE after affinity purification, cation exchange chromatography was applied followed by SEC to obtain a homogeneous population of the protein (Figure 4.1B and 4.1C). A single well-resolved peak was observed at an estimated molecular weight of 200 kDa and SDS-PAGE of fractions underneath the peak confirmed that the peak corresponded to Strep-tagged PlaB. SEC thus confirms that the purified protein is pure and monodisperse.

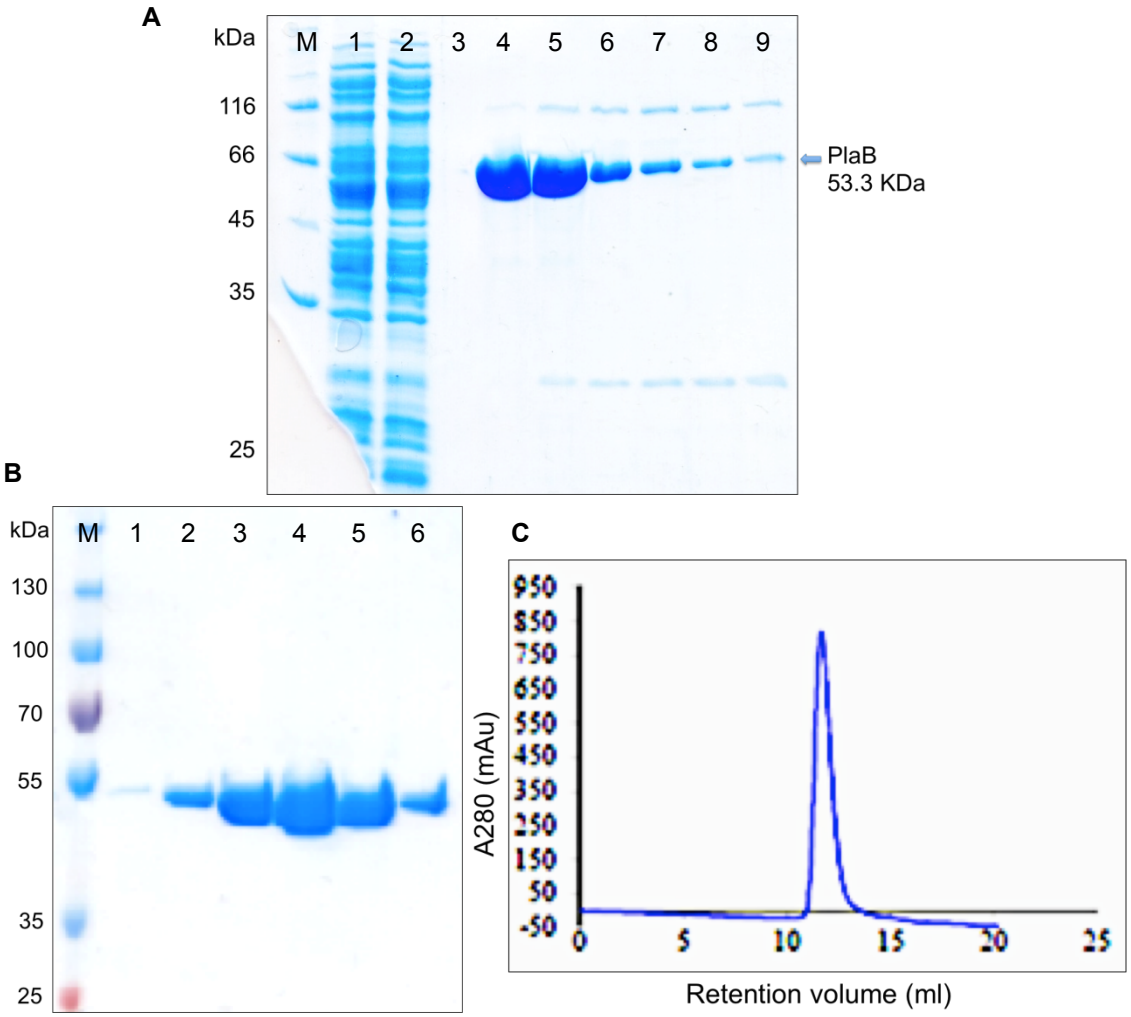


Figure 4.1: Analysis of purified PlaB. A: SDS-PAGE analysis Strep-tag affinity purification of PlaB at R.T. M: Protein marker; lane 1: flow through, lane 2-3: wash 1 and wash 2, respectively, lane 4-9: elution fractions. B: SDS-PAGE analysis of protein fractions eluted from gel filtration following cation exchange. M: protein marker; lane 1 – 6: gel filtration fractions. The protein amount loaded on SDS-PAGE was the same with the amount using for the SDS-PAGE analysis after affinity purification. Only PlaB band was observed and impurities were not detected. C: Gel filtration profile of PlaB at 12°C, showing single and relatively symmetrical peak. Gel filtration buffer: 100 mM HEPES pH 7.0, 100 mM NaCl; Flow rate: 0.3 ml/min; Sample volume: 0.5 ml.

However, after each step of chromatography a significant protein amount was lost. Eventually, 90% of protein amount was lost through steps of cation exchange, gel filtration and concentration. The obtained amount of protein, therefore, was only sufficient for initial crystallization screening but not for further crystallization optimization of the same batch. In order to obtain sufficient protein amount for crystallization optimization, the protein production was scaled up to 6-8 L of cell culture. At high yield, the protein was quickly precipitated (Figure 4.3A). Parameters that may cause protein precipitation or

enhance protein solubility were investigated and factors that were found to have clear contribution to the protein purification are described below.

4.1.3 Factors that have impact on protein purification

Among factors that were examined (including buffer compositions and pH, purification temperatures, salt types and concentrations, reducing agents), three factors were found to have clear effects on protein solubility that are temperature, the presence or absence of reducing agents, and the salt concentrations in protein buffers.

Temperature

Initially, the protein was purified at room temperature (R.T.). Although no precipitation was observed at small-scale purification, the sample could not be measured by DLS at 20°C, most likely because of high polydispersity that are out of range of DLS. When the measurements were performed at 4°C, however, it was possible to collect DLS data and the protein was relatively monodisperse (with polydispersity of 21.7%) (Figure 4.2). These results suggest that the protein was quickly aggregated at R.T and 4°C was necessary to keep the protein at monodisperse form. Thus, all steps of affinity protein purification were performed in cold room, all buffers were pre-chilled before using, the chromatography steps were performed using ÄKTA system in the fridge (at 12°C) and the collected fractions were moved to and stored in cold room as soon as possible before the purification procedure ends.

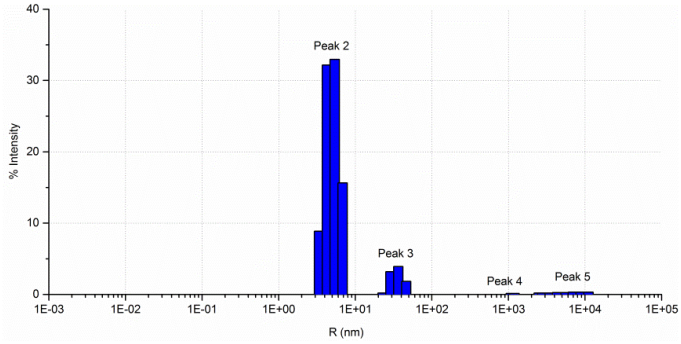


Figure 4.2: DLS analysis of purified PlaB. The first peak corresponds to the protein: At 4°C, the hydrodynamic radius of the protein is around 4 nm, with the percentage of polydispersity is 21.7, suggesting the protein is monodisperse. The other peaks may be aggregates or dust. Measurement at 20°C could not be obtained with the notification from the measurement software that the sample was out of measurement range of DLS.

Reducing agents

Although low purification temperature keeps the protein monodisperse during small-scale purification (from 1-2 L cell culture), it could not prevent protein precipitation at high yield (Figure 4.3A). The protein sequence of PlaB contains 5 cysteines. Information about disulfide bonding of the PlaB is currently unknown. The presence of cysteine gives high chance of non-specific cross-linking of free cysteines between protein molecules, especially at high concentration, resulting in protein aggregation and precipitation. On the other hand, as PlaB is expressed in host cytoplasm, which is in reducing condition, it is also possible that a reducing environment is required for PlaB stability. Because initial purification protocol did not contain reducing agent, a high concentration of DTT was added directly into precipitated protein solution to verify if precipitation had been caused by cysteines. The precipitated protein was reversed into soluble form (Figure 4.3B).

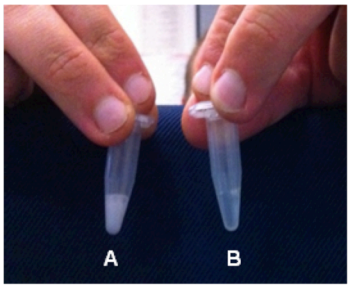


Figure 4.3: **Effect of reducing agent on protein precipitate.** A. Protein precipitated after large-scale purification. B: The protein precipitate was then treated by adding a small amount of 1 M DTT stock. The solution became transparent, indicating the protein was reversed into soluble form.

Salt concentration

Another factor that may affect solubility of the protein is salt concentration (Van Pham et al., 2010). A buffer screen by gel filtration in different NaCl concentration was thus conducted simultaneously. The results revealed that the protein tends to aggregate at high concentration of salt (Figure 4.4). It is consistent with the suggestion from ThermoFluor assay conducted by Kuhle (personal communication). In fact, protein precipitation was observed some times when the protein was eluted with high concentration into buffer that contains high concentration of salt during cation exchange. As a result, the yield of protein was significantly reduced or almost all of protein was lost once it was precipitated. This step however, is unavoidable for protein elution in cation exchange chromatography.

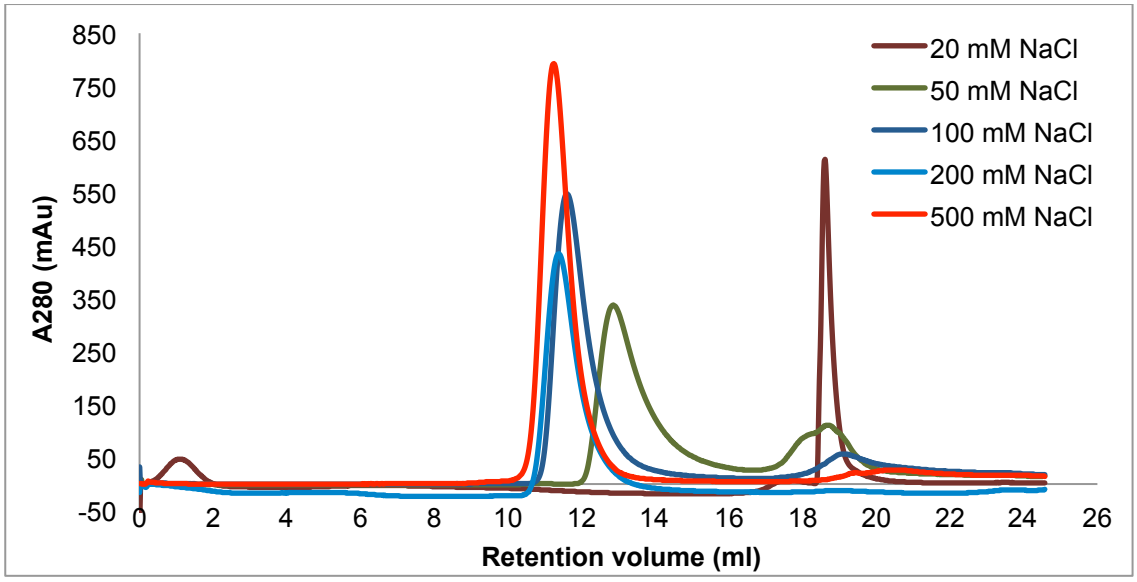


Figure 4.4: SEC of 4 mg protein injected in buffers at different NaCl concentrations. The protein was injected into a S200 10/300 column in running buffer consisting of 20 mM HEPES pH 7.0, 5 mM DTT, and NaCl at concentrations indicated above. Sample volume: 0.5 ml; Flow rate: 0.3 ml/min. As seen in the gel filtration profile of PlaB at each salt concentration, the higher the salt concentration, the earlier the protein eluted, except the one at 20 mM NaCl, which was eluted at much later retention volume. The other peak appeared in the gel filtration profile at 50 mM NaCl with much lower peak maximum was not PlaB and intractable on SDS-PAGE. They might be small molecules.

4.1.4 Reverse anion exchange with optimized buffers

A different protein purification scheme was therefore tested to overcome the effect of high salt concentration. Reverse anion exchange was applied after affinity purification. The PlaB protein was eluted during flow through of anion exchange and therefore was maintained in low salt (100 mM) buffer (Figure 4.5). 10 mg protein with high purity and homogeneity were obtained from 3 L of cell culture (Figure 4.6B). The protein can be flash-frozen with liquid Nitrogen and stored at -80° for long-term usage. There was no degradation sign after 2 weeks at 4°C and R.T. This final scheme is described in Figure 4.6C.

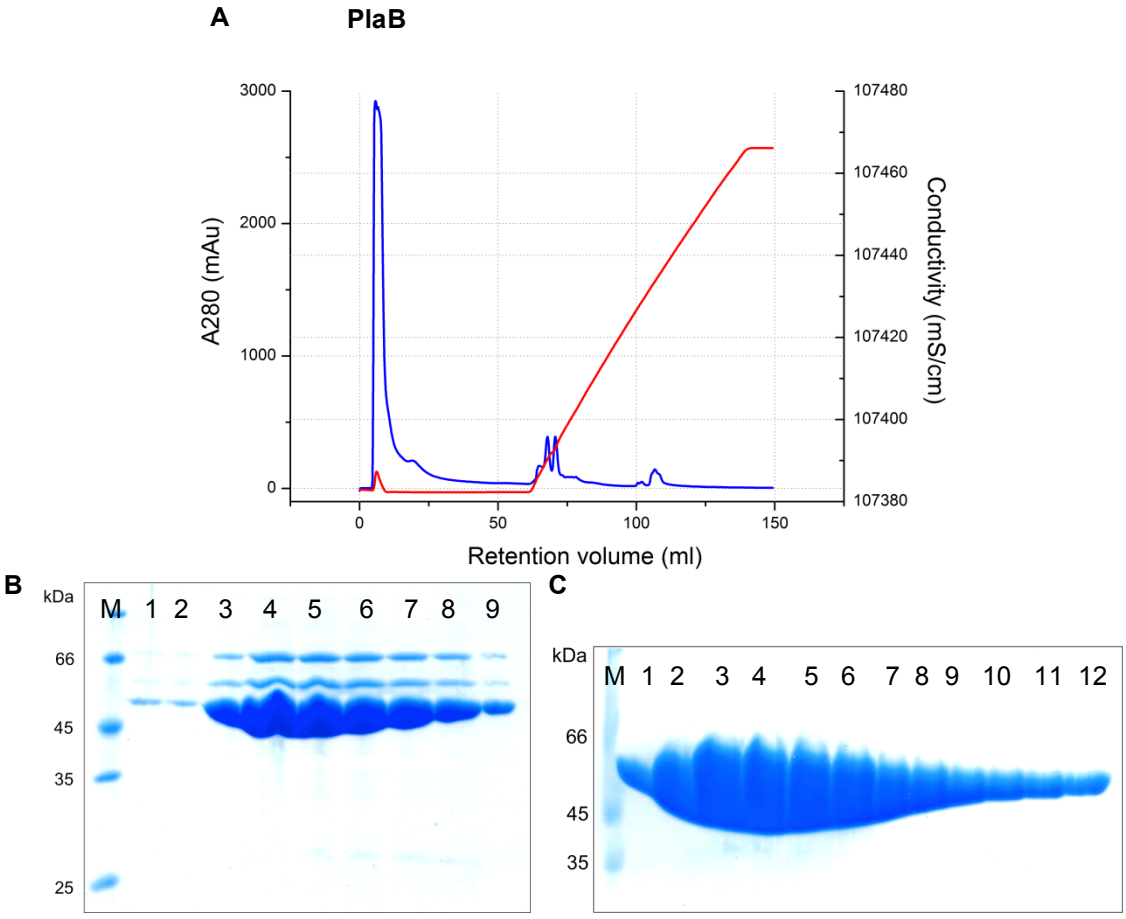


Figure 4.5: Reverse anion exchange purification of PlaB. A: Reverse anion exchange chromatogram of PlaB from a Mono Q column; blue: A280 of the protein; red: Conductivity. The protein did not bind to the column and could be collected in low-salt buffer (100 mM NaCl, low conductance), while impurities bound to the column and eluted later. B: SDS-PAGE analysis of the protein collected after affinity purification; lane M: protein marker, lanes 2-10: elution fractions. C: SDS-PAGE analysis of the protein collected after reverse anion exchange; lane M: protein marker, lanes 1-12: collected fractions. The collected fractions were intentionally loaded at high amount to access the protein purity. The impurities were not seen on Coomassie stained SDS-PAGE even at doubled protein amount loaded on the gel.

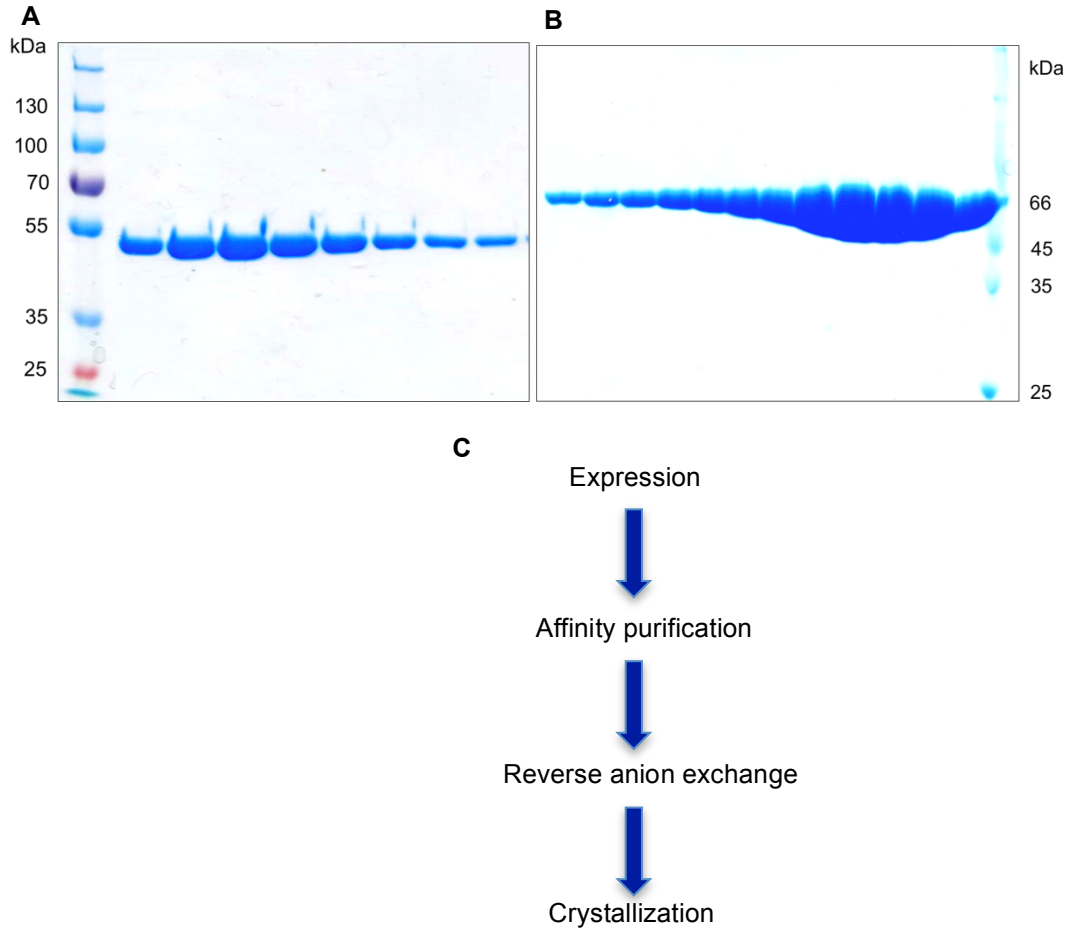


Figure 4.6: Comparison of protein yield purified from two above-described schemes. A: Protein purified from the first scheme. In this scheme scheme, the protein was purified through three steps: affinity purification, cation ion exchange and SEC. The protein was then pooled down and concentrated to 5 mg/ml. The final amount of protein is 2 mg from 3 L of cell culture. B: Protein purified from the optimized scheme. In this scheme, reverse anion exchange was following affinity purification as the final purification step. The final amount of protein after being concentrated is 10 mg from 3 L of cell culture. C: Final protein preparation scheme: The protein is expressed in Tuner cells at 20^0C overnight. The cell pellet was harvested and purified in two steps, Strep-tag affinity purification and reverse anion exchange. The protein is then concentrated to either 5 mg/ml or 10 mg/ml, aliquoted and flash frozen by liquid Nitrogen for long-term storage at -80^0C. There is no change in crystallization behavior of the protein before and after being frozen.

4.1.5 Simplifying protein preparation procedure

In the first year of the project, the protein was prepared by 2 persons at 2 places. First the protein expression at small scale and affinity chromatography purification was carried out in Robert Koch Institute. The protein was then either picked up or brought over to the HZI for further steps of purification and crystallization.

As mentioned above, a large amount of purified protein in the same batch was required. Due to the lack of capacity for large-scale protein expression in Robert Koch Institute, large-scale protein expression was moved to the HZI. However, as wild type protein has hemolytic activity, it is not allowed to work with the protein in S1-lab and an S2-lab was not accessible at that time. Hence, the inactive PlaB mutant D203N was used. No change in physical behaviors of D203N mutant compared to the wild type was observed.

4.2 Characterization of PlaB at different concentrations

4.2.1 Size-exclusion chromatography (SEC) analysis

For SEC analysis, the protein was injected into S200 10/300 column in buffer consisting of 20 mM HEPES pH 7.0, 100 mM NaCl, 5 mM DTT. A significant shift of the retention volume of the protein at low and high concentration (Table 4.1, Figure 4.7) has been observed. While the retention volume of the protein at 1 mg/ml was 13.54 ml, the eluted peak of the protein at 10 mg/ml was 12.51 ml, which raises the question whether PlaB changes its oligomerization state depending on the protein concentration.

Table 4.1: Retention volumes of PlaB in SEC

Protein	Known MW (kDa)	Concentration (mg/ml)	Retention volume of peak maxium (ml)
Conalbumin	75	1	14.2
PlaB		1	13.54
Aldolase	158	1	12.75
PlaB		10	12.51

The retention volumes were determined by integrating the A280 absorbance peak of SEC using UNICORN software.

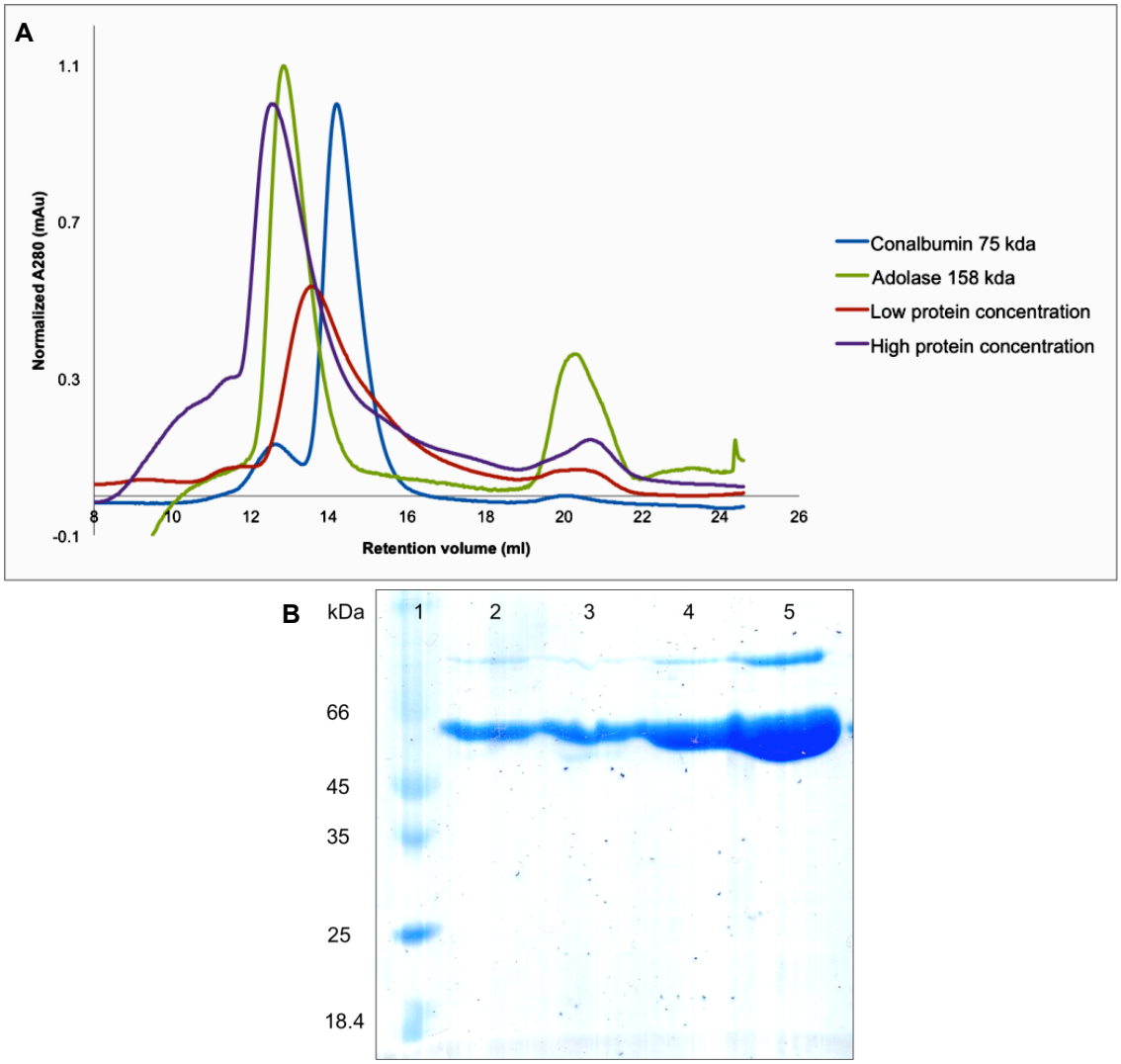


Figure 4.7: SEC analysis of PlaB at low and high protein concentrations. A: SEC profiles of tested proteins. All proteins were injected into a S200 10/300 column in running buffer consisting of 20 mM HEPES pH 7.0, 100 mM NaCl and 5 mM DTT. Sample volume: 0.5 ml; Flow rate: 0.3 ml/min. The absorbance 280 nm was normalized to 1. B: SDS-PAGE analysis of PlaB collected from the SEC run; lane 1: protein marker, lane 2-3: fractions collected from highest gel-filtration peak of PlaB at the concentration of 1 mg/ml, lane 4-5: fractions collected from gel-filtration peak of PlaB at the concentration of 10 mg/ml.

To estimate the oligomerization state of the PlaB at these two concentrations, Aldolase and Conalbumin were used as controls. A linear function (y = a + bx) showing correlation between logarithm of molecular weight and corresponding retention volume was established based on known molecular weights of Adolase and Conalbumin (Whitaker 1963). The molecular weight of PlaB monomer is 53 kDa. The results suggested that 13.54 ml corresponds to a molecular weight of 116.2 kDa, which is approximately a dimer, and 12.51 ml corresponds to a molecular weight of 195.79 kDa, which is approximately a

tetramer (Table 4.2). The results are consistent with results from small-angle X-ray scattering, which suggested that PlaB is a tetramer and probably composed of a dimer of dimers (Kuhle, personal communication).

Table 4.2: Calculation of MW of PlaB at low and high protein concentrations

Protein	Known MW (kDa)	a	b	x (ml)	y = a+ bx = log(MW)	Calculated MW (kDa)
Conalbumin	75.00	5.04	- 0.22	14.20	1.92	
Adolase	158.00	5.04	- 0.22	12.75	2.24	
PlaB (1 mg/ml)		5.04	- 0.22	13.54	2.07	116.20
PlaB (10 mg/ml)		5.04	- 0.22	12.51	2.29	195.79

4.2.2 Dynamic Light Scattering (DLS) analysis

DLS is usually used to evaluate the aggregation state and to measure polydispersity of a protein (Proteau et al. 2010). Percentage of polydispersity is predictive of crystallisability (Borgstahl 2007). It is not suggested to use DLS to determine molecular weight of the protein, but it is possible to observe if there is any change in oligomerization state of the protein at different concentrations.

For DLS experiment, the protein was measured at 4 different concentrations: 1 mg/ml, 5 mg/ml, 7.5 mg/ml and 10 mg/ml. At three latter concentrations, there is no significant change in R_H of the protein, suggesting that there is no change in oligomerization state at the protein concentration from 5 mg/ml to 10 mg/ml (Figure 4.8). At 5 mg/ml and 7.5 mg/ml, the protein was monodisperse, suggesting these concentrations are suitable for crystallization. It was not possible to obtain a qualified reading with a protein concentration of 1 mg/ml due to increased baseline and very high SOS error, suggesting that at low concentration, the protein is highly polydisperse.

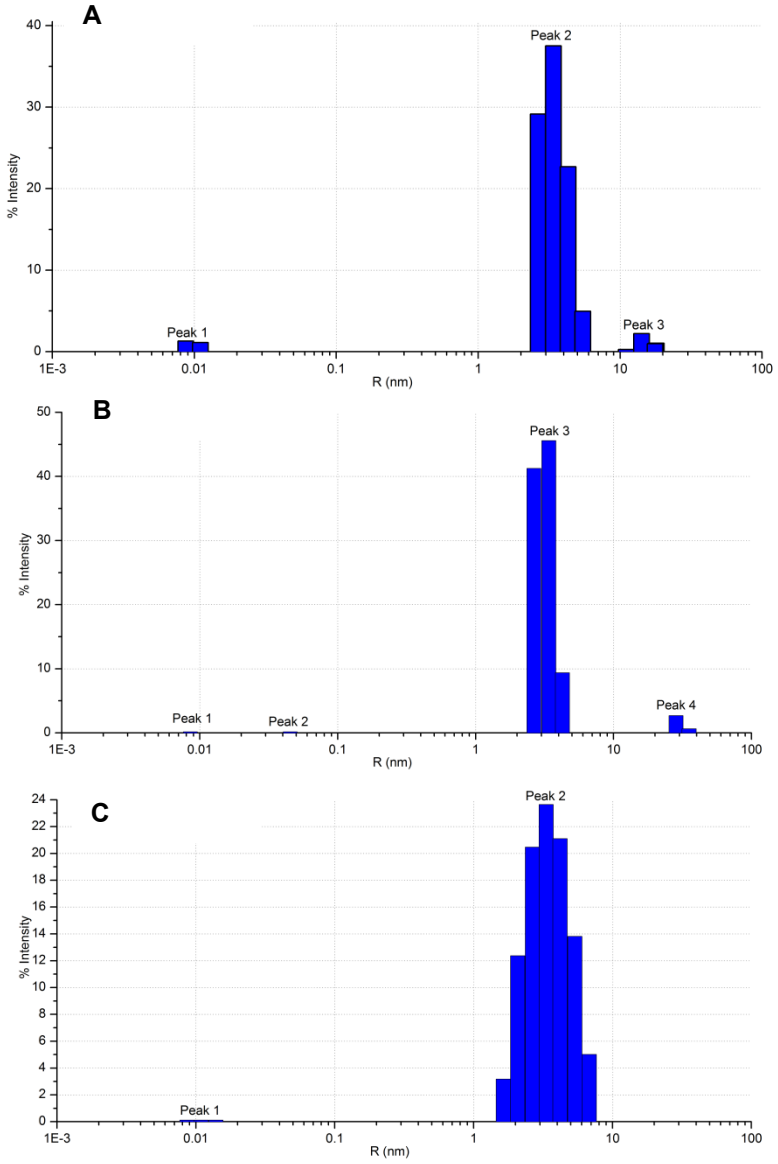


Figure 4.8: DLS profile at different concentrations after purification. 25 measurements were carried out for each sample at 4°C. A. Protein concentration of 5 mg/ml (R_H = 3.5 nm; %PD = 21.7); B. Protein concentration of 7.5 mg/ml (R_H = 3.1 nm; %PD = 15.9); C. Protein concentration of 10 mg/ml (R_H = 3.6 nm; %PD = 35.6). The %PD equal to or smaller than 20% indicates the sample was monodisperse and R_H of about 3 – 3.5 nm is in the range of non-aggregated protein molecules.

4.3 Crystallization

4.3.1 Initial crystallization

Based on the DLS results, purified protein was screened at the concentrations of 5 mg/ml against various commercial crystallization kits covering 960 different

conditions. Right before crystallization, the protein solution was centrifuged at 14,000 rpm at 4°C for 20 minutes to remove any possible precipitation. The crystallization screening was carried out at R.T in a sitting drop setup and the crystallization plates were incubated at 4°C and 20°C. Three crystal forms were obtained (Figure 4.9) in a broad range of screening conditions. PEG was the precipitant in most of these conditions. Crystals of suitable size were subjected to x-ray diffraction experiments but the diffraction resolution was not better than 8 Å. Conditions that yielded best diffracted crystals (Figure 4.9A & B) were chosen for further optimization.

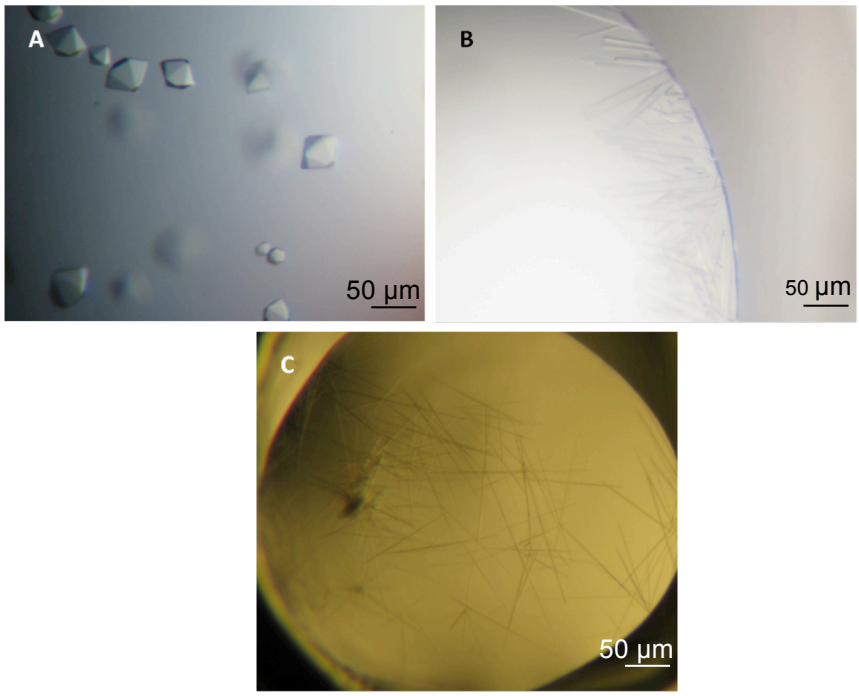


Figure 4.9: Initial PlaB crystals. Crystals appeared after mixing 0.2 µl of the protein solution at 5 mg/ml with 0.2 µl of the according reservoir followed by incubation in a sitting drop setup at 20°C. A: hexagonal crystals obtained from 0.1 M MES pH 6.5, 15% PEG 400; B: rod-shaped crystals obtained from 0.1 MES pH 6.5, 10% PEG 6000; C: An example of needles appeared in various conditions.

4.3.2 Triclinic crystals

Rod-shaped crystals appeared in several initial conditions but the condition in which crystals showed diffraction was 0.1 M MES pH 6.5, 10% PEG 6000. PEG concentration screen in the range of pH from 5.5 to 8.5 gave the optimized crystals in 0.1 M MES pH 5.5, 5% PEG 6000 at 20°C. The best crystal diffracted to a resolution of 2.6 Å at a micro-focus beamline ID-13 at ESRF

(Figure 4.10). Additive screen was also tried in the attempt to obtain bigger and thicker crystals. However none of the additives improved the crystal quality. Several cryoprotectants were tested including glycerol, PEG 400, PEG 200, and paraffin. The crystals diffracted best and no ice ring was seen when using 20% glycerol as the cryoprotectant. A solution of reservoir buffer including 20% glycerol was prepared, the crystal transferred from its drop into this solution and then quickly frozen in liquid nitrogen.

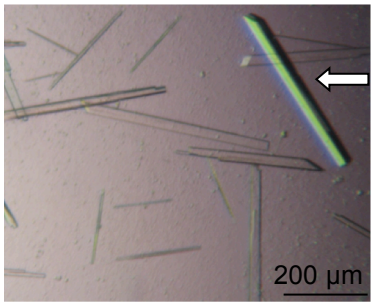


Figure 4.10: Optimized triclinic crystals. Optimized crystal obtained after mixing 1 µl of the protein solution at 5 mg/ml with 1 µl of the according reservoir followed by incubation in a hanging drop setup at 20°C. The crystal was quickly dipped into cryoprotectant solution and then frozen in liquid nitrogen. X-ray measurement of this crystal was carried out at microfocus beamline ID-13 at ESRF.

4.3.3 Hexagonal crystals

The initial crystallization condition of hexagonal crystals was 0.1 M MES pH 6.5, 15% PEG 400. The crystal diffracted to a resolution of 5 Å at BESSY synchrotron. To improve the diffraction resolution, the condition was optimized by focusing on the precipitant and the pH. A screen of different PEGs with different molecular weights gave a crystal with improved diffracting resolution of 2.5 Å.

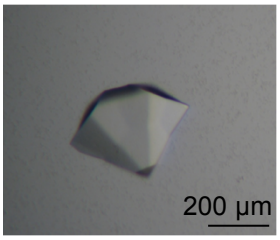


Figure 4.11: Optimized hexagonal crystals. Optimized crystals obtained after mixing 1 µl of the protein solution at 5 mg/ml with 1 µl of the according reservoir followed by incubation in a hanging drop setup at 20°C. The crystal was quickly deep into cryoprotectant solution and then frozen in liquid nitrogen. X-ray measurement of this crystal was carried out at beamline ID-29 at ESRF.

4.3.4 Significant observations during protein crystallization

Three observations made during the protein purification process contributed to the optimization of protein crystallization, and vice versa. Firstly, the protein behavior was irreproducible among different batches of protein purification (Figure 4.12). This further emphasized the need of preparing protein in large scale for higher reproducibility of crystallization (See section 4.1.2).

Secondly, crystallization of the protein in the presence of DTT or TCEP as an additive suggested that DTT was better for protein crystallization. The crystals were able to form in the presence of DTT but not TCEP. 5 mM DTT, therefore, was then added into buffers at every steps of protein preparation.

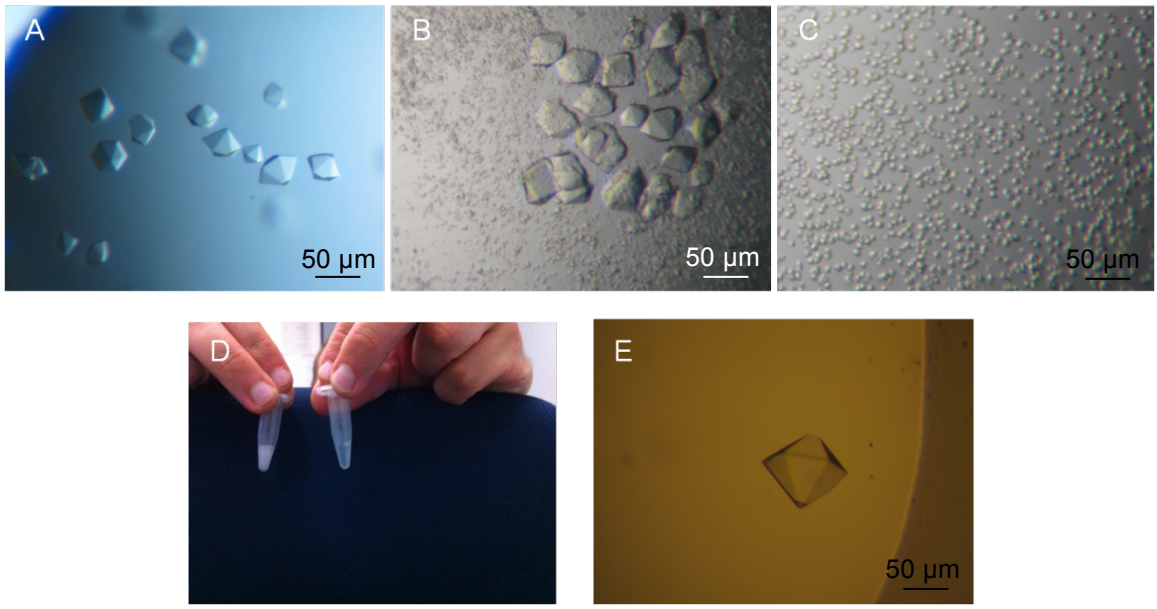


Figure 4.12: PlaB crystal's behavior among batches. A-C: Crystals of three different batches of protein preparation at small scale (from 1-2 L cell culture). D: Large scale protein preparation in buffer without DTT resulted in aggregation (See section 4.1.3). E: Crystal obtained from a large scale protein preparation batch in buffer containing DTT.

Thirdly, the crystals were obtained at 20°C. An SDS-PAGE analysis showed that there is no visible degradation of the protein in the crystals after two weeks (Figure 4.13). As mentioned in section 4.1.4, there was also no sign of protein degradation in solution. No degradation at this temperature gives good chance for the protein to be crystallized in full-length form.

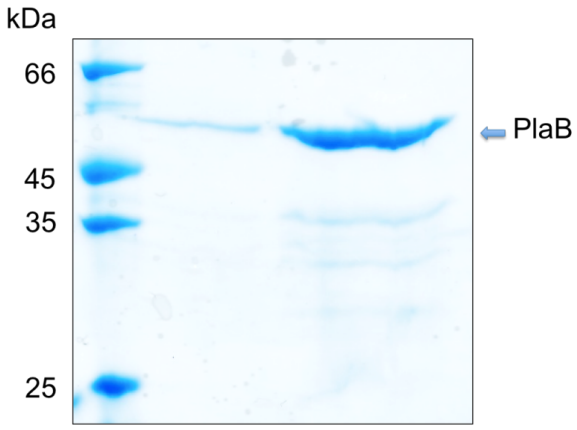


Figure 4.13: SDS-PAGE analysis of PlaB crystals. From left to right: lane 1: protein marker; lane 2: PlaB crystals dissolved in sample buffer; lane 3: a different batch of purified PlaB

4.4 Crystallographic analysis

Datasets were collected at ESRF micro-focus beam line for rod-shaped crystals and at BESSY for hexagonal crystals. The rod-shaped crystal diffracted up to 2.6 Å and belongs to P1 space group. The hexagonal one diffracted up to 2.7 Å and belongs to $P6_x22$ space group. The datasets were processed with XDS (Kabsch 2010). The numbers of monomers per asymmetric unit were calculated using Mathews cell content analysis in *CCP4* suite (Diseases 1968; Winn et al. 2011; Kantardjieff & Rupp 2003) (Table 4.3).

Processing of datasets of hexagonal crystal using L-test in *CCP4 suite* (Padilla & Yeates 2003; Winn et al. 2011), however, revealed that the crystal is a perfect twin.

Table 4.3: Data statistics of the optimized crystals

	The hexagonal crystal	The rod-shaped crystal
Resolution (Å)	20.0 - 2.7 (2.8 - 2.7)	20 - 2.6 (2.7 – 2.6)
Space group	$P6_X22$	P1
Average unit cell dimensions (Å)	197.9 / 197.9 / 175.2	102.1, 89.7, 104.5
α /β/γ (°)	90.0 / 90.0 / 120.0	75.9, 90.1, 93.6
Rmerge (%)	9.4 (75.4)	9.3 (43.4)
Mean ((I)/σ(I))	15.4 (2.4)	7.5 (1.8)
Completeness (%)	98.4 (97.9)	92.7 (84.3)
No. of monomers per au	4	8

Additive screen and crystallization at high temperature (30°C) were tested in the attempt to solve twinning problem. Expected untwinned dataset has L statistics of 0.5 while L statistic of a perfect twin is 0.375. Interestingly, the crystal obtained with the presence of potassium chloride as an additive has L statistic increasing to 0.438, near the value of an untwinned dataset. Some crystallization conditions producing rod-shaped crystals also contained potassium chloride as a salt, suggesting that potassium chloride may play a role in crystal packing, and further experiments with different concentrations of potassium chloride are worth to be tried.

With a perfect twin, twinned intensities can increase 50% more and thus diminishes the inherently small anomalous signal (which only add 0.5% to 15% to the diffraction intensity). Experimental phasing has been rarely successful with twinned data (Rupp 2010). Molecular replacement solutions are in general less sensitive to twinning, and in these cases the problem usually becomes evident in the refinement stage. However current progress in computer program allows refinement against twinned data, made molecular replacement possible to use for phasing with twinned data. Therefore, molecular replacement was tested for the twinned dataset of PlaB.

4.5 Molecular replacement

Swiss-Model (Arnold et al. 2006), FUGUE (Shi et al. 2001), NCBI protein blast and Phyre (Kelley & Sternberg 2009) were used to search for structural homologs of PlaB. Among these programs, Phyre gave models of known structures with highest sequence identity and appropriate sequence alignment. Except the first template with unknown function, the others also belong to lipases (Figure 4.14). The highest sequence identity, however, is only 21% and all the templates only cover the N-terminal part of the protein. The highest coverage is only 37%. In order for molecular replacement to be successful, the sequence identity must be around 30% and the highest coverage of the search model must be at least 50% (Evans & McCoy). Sequence alignment between the model and PlaB suggests similarity in secondary structural contents at N-terminal part of PlaB. However, it was only able to align the two residues Ser 85 and Asp 203 with the respective ones in the search model. The third residue, His 251 is out of alignment area (Figure 4.15). Given all these information, no templates found were qualified as search model for molecular replacement. Attempts were then put on experimental phasing of triclinic crystals. The conventional heavy-atom derivatization method was first pursued.

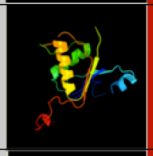
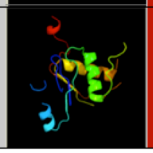
#	Template	Alignment Coverage	3D Model	Confidence	% i.d.	Template Information
1	c3ds8A	Alignment		99.8	15	**PDB header**: structural genomics, unknown function **Chain**: A: **PBD Molecule**:lin2722 protein; **PDBTitle**: the crystal structure of the gene lin2722 products from listeria2 innocua
2	d1tcaa	Alignment		99.8	14	**Fold**: alpha/beta-hydrolases **Superfamily**: alpha/beta-Hydrolases **Family**: Fungal lipases
3	d1ku0a	Alignment		99.8	20	**Fold**: alpha/beta-hydrolases **Superfamily**: alpha/beta-Hydrolases **Family**: Bacterial lipase
4	d1ji3a	Alignment		99.7	21	**Fold**: alpha/beta-hydrolases **Superfamily**: alpha/beta-Hydrolases **Family**: Bacterial lipase
5	c2hihB	Alignment		99.7	17	**PDB header**: hydrolase **Chain**: B: **PBD Molecule**: lipase 46 kda form; **PDBTitle**: crystal structure of straphylococus hyicus lipase

Figure 4.14: Potential structural homologs of PlaB detected by Phyre II (Kelley & Sternberg 2009). The first column is the PDB id of the template. Alignment coverage indicates the area of sequence alignment between the models and PlaB.

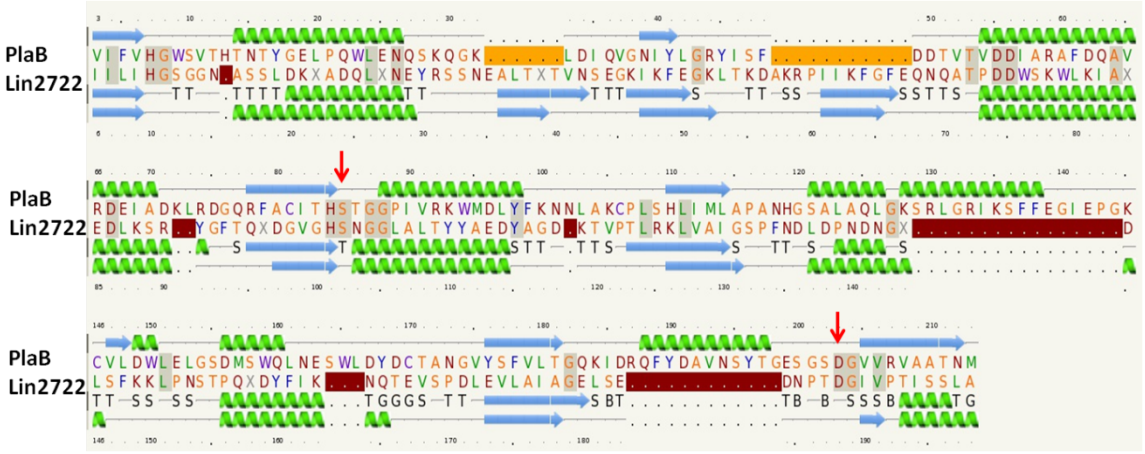


Figure 4.15: Sequence and secondary structure alignment of search model (Lin2722) and PlaB. Red arrows indicate catalytic residues.

4.6 Heavy atom screening and co-crystallization

HATODAS 2 server was used to predict which heavy atoms bind to the PlaB (Sugahara et al. 2009) . The result suggests that Sm, Hg, Au and Pt derivatives

were the most potential candidates. These derivatives were then screened by native gel to rule out the ones that do not bind to the protein or denature the protein (Boggon & Shapiro 2000). It is assumed that a heavy atom binding to the protein will lead to a shift on native gel because it provides additional charges to the protein. And if a protein is denatured by a heavy atom, it would precipitate and cannot move into the well of the native gel, resulting in the empty corresponding lane on native gel. An optimized protocol for native gel preparation and electrophoresis was established (Table 4.3). Figure 4.16 shows that the protein was denatured in the presence of KAuBr, $HgCl_2$, $Hg(CH_3COO)_2$ and K_2PtCl_6 and did not bind to Sm derivatives. The pattern in lane 6 and lane 7 are different from the one in lane 1. It is known that mercury binds to sulfur of sulfhydryl group of cysteine [Ref], and because there are 5 cysteins in the protein sequence, it is highly possible that $C_3H_6HgO_2$ and $C_2H_7HgO_4P$ might have bound to the protein. $C_3H_6HgO_2$ and $C_2H_7HgO_4P$ were thus chosen for further experiments. I3C and I4C were also tested.

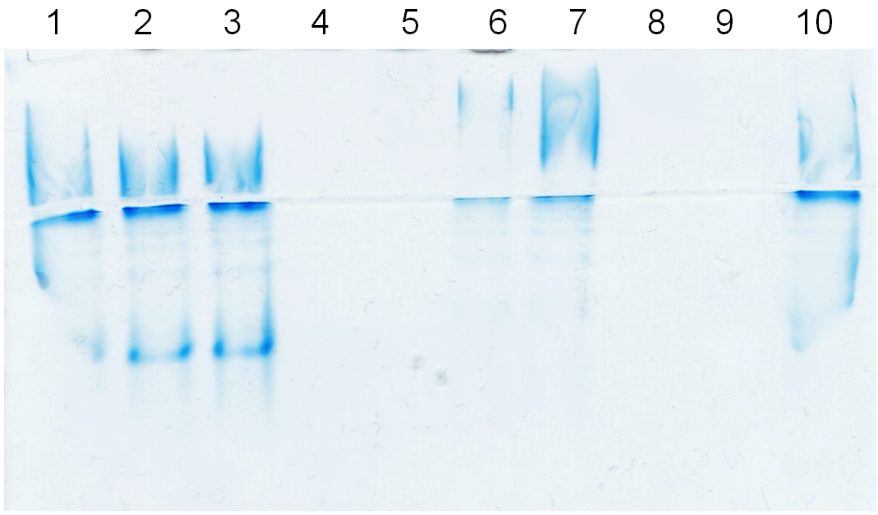


Figure 4.16: Heavy atom screening. Lane 1: Native protein; Lane 2, 3, 10: Sm derivatives; Lane 4: KAuBr, Lane 5: $Hg(CH_3COO)_2$;Lane 6: $C_3H_6HgO_2$; Lane 7: $C_2H_7HgO_4P$; Lane 8: $HgCl_2$; Lane 9: K_2PtCl_6. Lane 4, 5, 8, 9 were blank, an indication that the protein was severely denatured once mixing with respective heavy atoms, and precipitated, thus could not move into the wells. The pattern of lane 2, 3 and 10 are not different from lane 1, suggesting the Sm derivatives did not bind to the protein. Only mixtures running in lane 6 & 7 produced a relatively different pattern compared with the control in lane 1. Thus, $C_3H_6HgO_2$ and $C_2H_7HgO_4P$ were chosen for preparing derivative crystals.

To increase the opportunity for uniformity and completeness of heavy atom binding, and because the triclinic crystals were very fragile, they were co-

crystallized with the protein at both conditions that gave hexagonal and triclinic crystals. No triclinic crystals were obtained. Co-crystallization of the protein with $C_3H_6HgO_2$ resulted in crystals with bigger size than the native one of the same batch, but the size of each derivative crystal was still below 50 µm and the crystals had no diffraction (Figure 4.17). I3C and I4C did not give crystals with sufficient size for measurement (Figure 4.18).

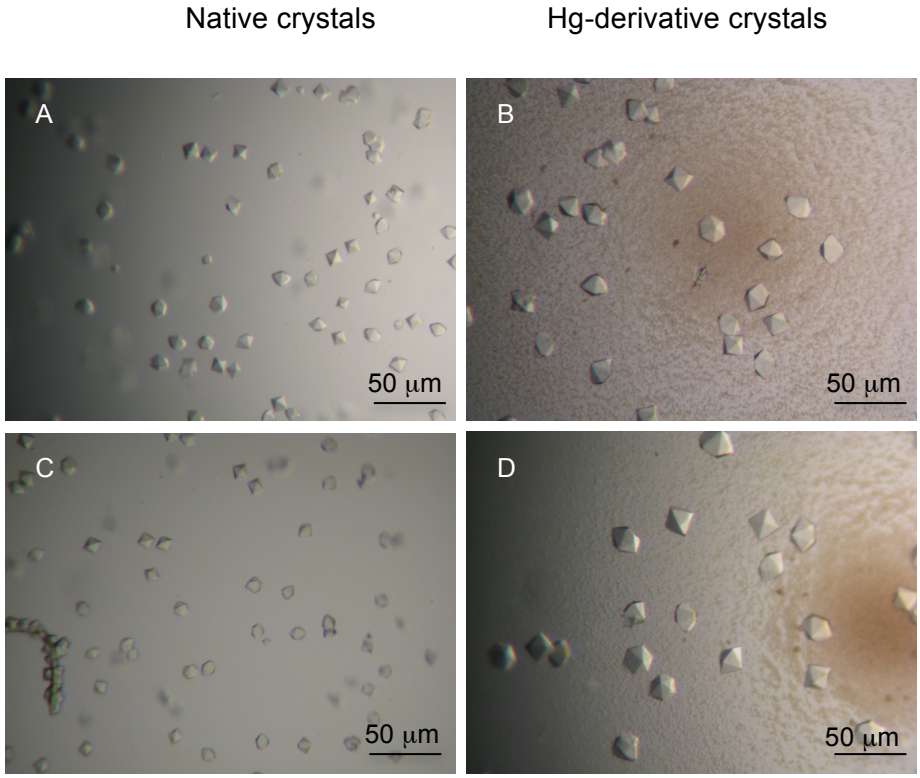


Figure 4.17: Hg-derivative crystals

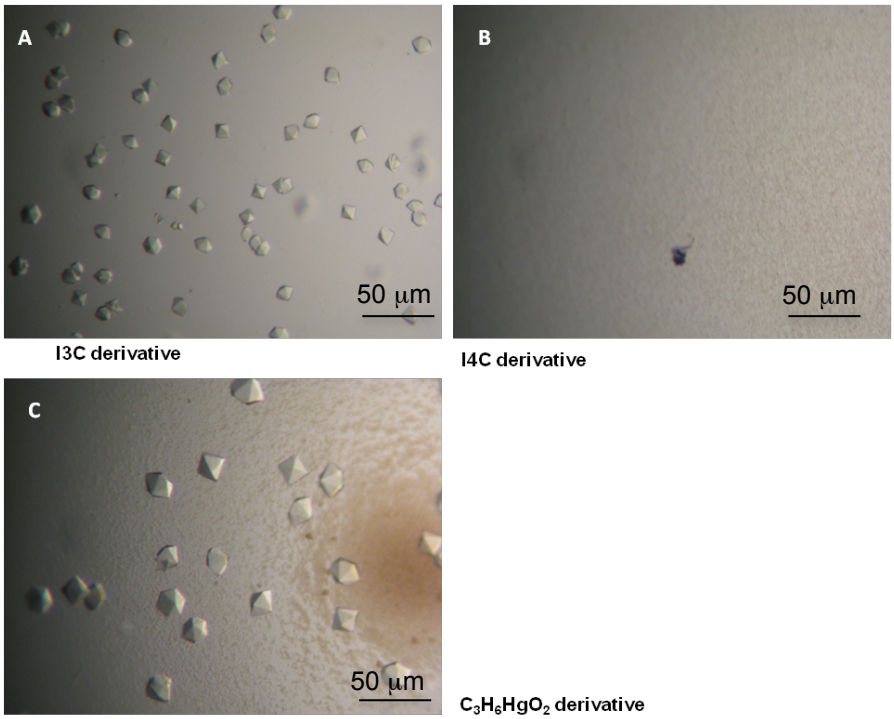


Figure 4.18: Different derivative crystals. A: I3C derivative; B: Co-crystallization of PlaB with I4C did not give crystals. C: $C_3H_6HgO_2$ derivative.

The protein concentration that was used for crystallization was 5 mg/ml. One possible reason why $C_3H_6HgO_2$ crystal has bigger size than the native one is that: in the presence of heavy atom, part of the protein was denatured, resulted in lower concentration of the protein and lower nucleation rate during crystallization process. To test this hypothesis, different concentrations of the protein were tested. However, there were no crystals produced at the protein concentration lower than 4 mg/ml. A protein concentration at 7.5 mg/ml and 10 mg/ml resulted in quick aggregation or shower of needles.

4.7 Establishment of an expression scheme for SeMet-labeled protein

Because heavy atoms did not give a promising result, SeMet was considered for production of derivative protein. There are two main stages in expression of SeMet-labeled protein. The first step was to grow the cells to high density (OD_{600} = 4). The cells were then harvested, washed and then divided into 4 parts. In the second step, the cells were diluted into new media so that OD_{600} was kept at 1. The cells were re-grown and induced in the presence of SeMet.

Four different expression schemes were tested (Table 4.4). In the first scheme, high density of cells were attempted to obtain by growing cells in minimal media plus methionine. The cells were then re-grown in minimal media in the presence of SeMet. In the second scheme, a mixture of amino acids (K, I, L, V) was added in the media at both steps. In the third scheme, minimal media was replaced by LB in the first step. No amino acids were added in both steps except SeMet was added in the second step. These three schemes of expression, however, resulted in cell death during growth or expression.

Table 4.4: Schemes that were tested for expression of SeMet-labeled protein

Scheme	Conditions		Result
	Stage 1	Stage 2	
1	Minimal media + Met	Minimal media + SeMet	Failed
2	Minimal media + amino acid mix	Minimal media + amino acid mix + SeMet	Failed
3	LB medium	Minimal media + SeMet	Failed
4	LB medium	Minimal media + amino acid mix + SeMet	Successful

To make the healthiest environment for cell growth, a protocol for expression of labeled-protein prepared for NMR was adapted. In the first step, LB was used to boost cell growth. In the second step when minimal media must be used to inhibit the synthesis of native methionine, the mixture of amino acids was added as supplements. With this protocol, it was able to keep the cell growing normal during induction. It was able to obtain 1 mg purified protein from 4 L of cell culture ready for crystallization (Figure 4.19). The incorporation rate of SeMet was nearly 90%.

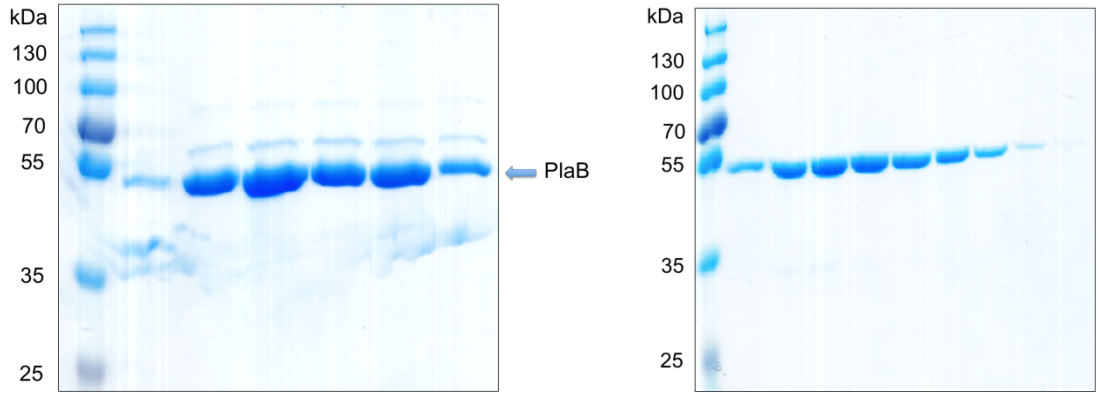


Figure 4.19: SDS-PAGE analysis of purified SeMet protein. A: After affinity purification. From left to right: lane 1: marker, lane 2-7: elution fractions of affinity purification. B: After reverse anion exchange. From left to right: lane 1: marker, lane 2-9: fractions collected from reverse anion exchange. The band on SDS-PAGE was confirmed by Mass spec that it belongs to PlaB protein.

4.8 Crystallization and crystallographic analysis of SeMet-labeled protein

The best crystals after several rounds of optimization were obtained in the condition consisting of 0.1 M MES pH 6.5, 15% PEG 6000 or 0.1 M Tris pH 8.5, 5% PEG 6000 and 50 mM KCl (Figure 4.20). The crystal could diffract up to 2.29 Å under in-house beam line. The SeMet-labeled protein, however, was quickly aggregated just within 2 weeks even in -80°C storage.

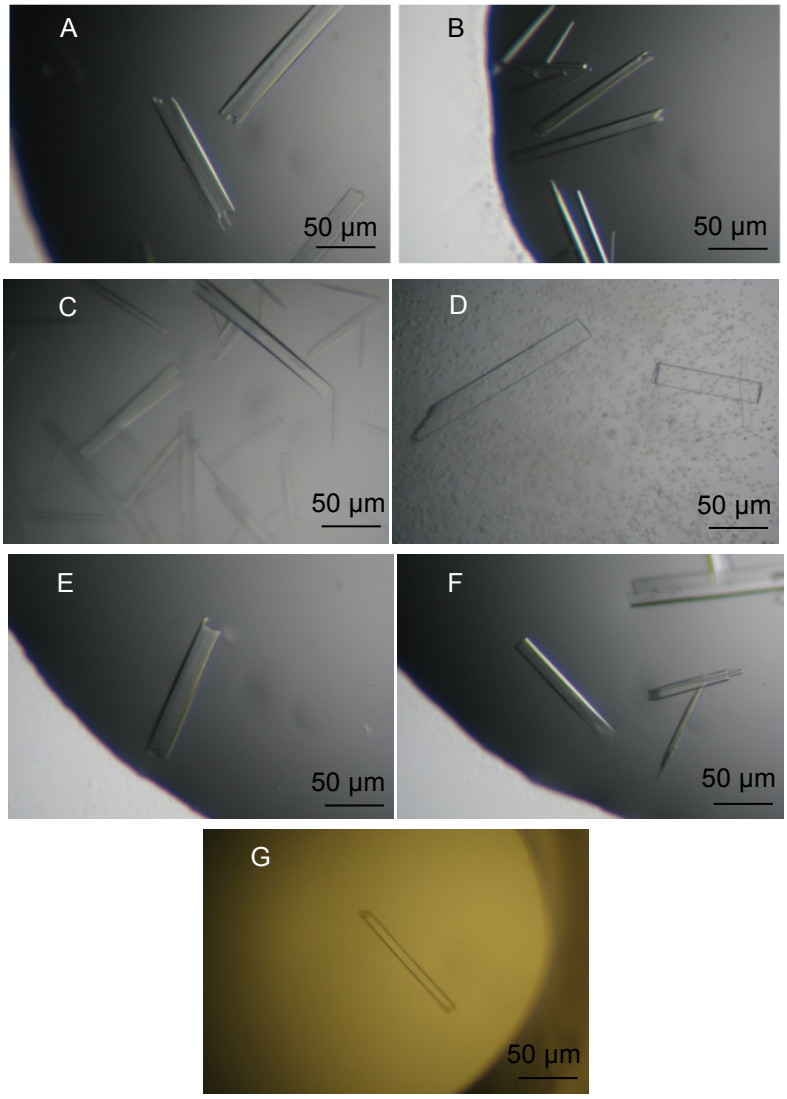


Figure 4.20: SeMet optimized crystals. A: 0.1 M MES pH 6.5, 15% PEG 6000; B: 0.1 M Tris pH 8.5, 50 mM KCl, 5% PEG 6000; C: 0.1 M Tris pH 7.0; 3% PEG 6000, 0.1 M Potassium chloride; D: Tris pH 8.0, 3% PEG 6000, 0.1 M Potassium chloride; E: Tris pH 8.5, 3% PEG 6000, 0.1 M Potassium chloride ; F: Tris pH 9.0, 3% PEG 6000, 0.1 M Potassium chloride; G: 0.2 M MgCl2, 0.1 M HEPES pH 7.5, 15% (w/v) PEG 400.

Se-absorbance scanning indicated that SeMet was incorporated into the crystal (Figure 4.21).

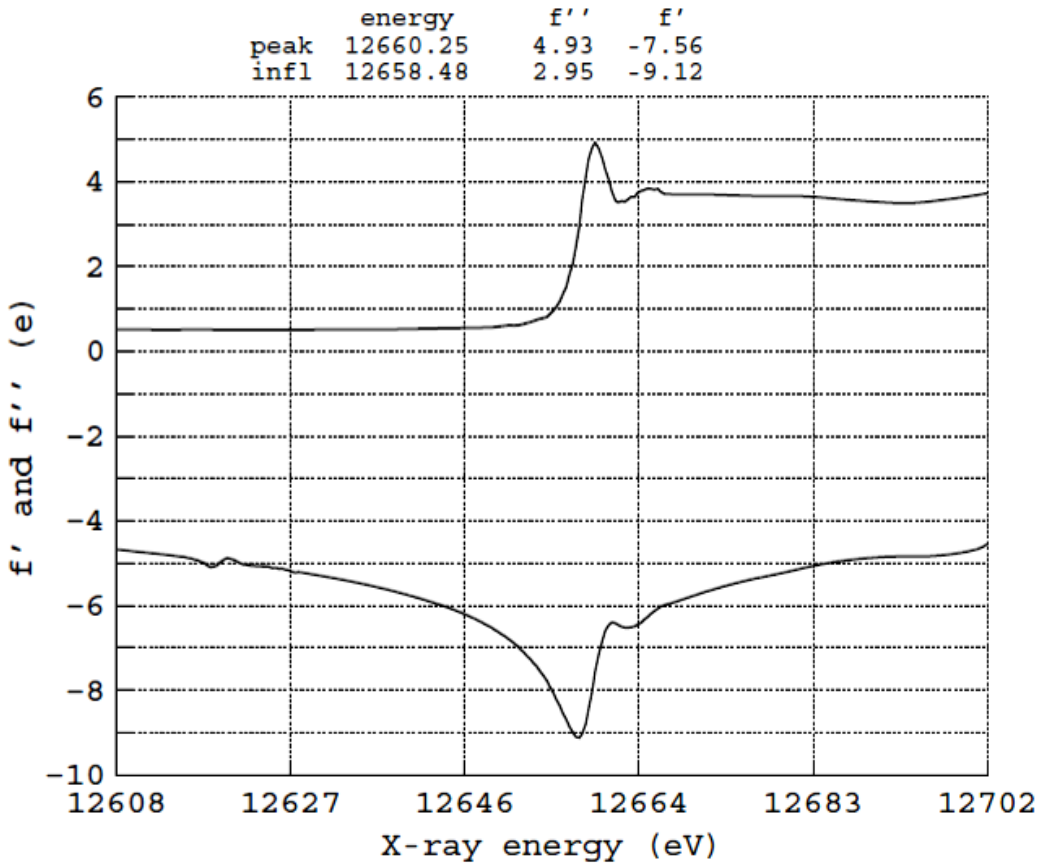


Figure 4.21: Se absorbance scan. F' and f'' are real and imaginary components of anomalous scattering factor, respectively. Typical anomalous scattering factors, f''max and f'min, estimated from X-ray spectra of protein crystals taken at MAD experimental stations, are 6 and -11, respectively (Wu et al. 1994).

Data statistics of SeMet-labeled crystals indicated that the derivative crystals belong to the same space group with the native triclinic one, with almost the same unit dimensions (Table 4.5). However, an analysis of diffraction pattern of the crystal (Figure 4.22), together with Rmeas analysis versus frames (Figure 4.23) revealed that the crystal has severe anisotropy, made it unable to be used for phase determination.

Table 4.5: Data statistics of SeMet crystals

	Native	SeMet derivative
Resolution (Å)	20-2.6 (2.7-2.6)	20 - 3.1 (3.2 – 3.1)
Space group	P1	P1
Unit cell dimensions		
a, b, c (Å)	75.9, 90.0, 93.6	75.9, 89.7, 93.3
α, β, γ (°)	102.1, 89.7, 104.5	102.4, 90.0, 104.2
Rmerge (%)	9.3 (43.4)	28.8 (80.0)
Mean I/σ (I)	7.5 (1.8)	6.9 (2.7)
Completeness (%)	92.7 (84.3)	98.6 (98.3)
Redundancy	2.0	5.5 (5.5)
SigAno > 1.1		20 Å - 3.9 Å

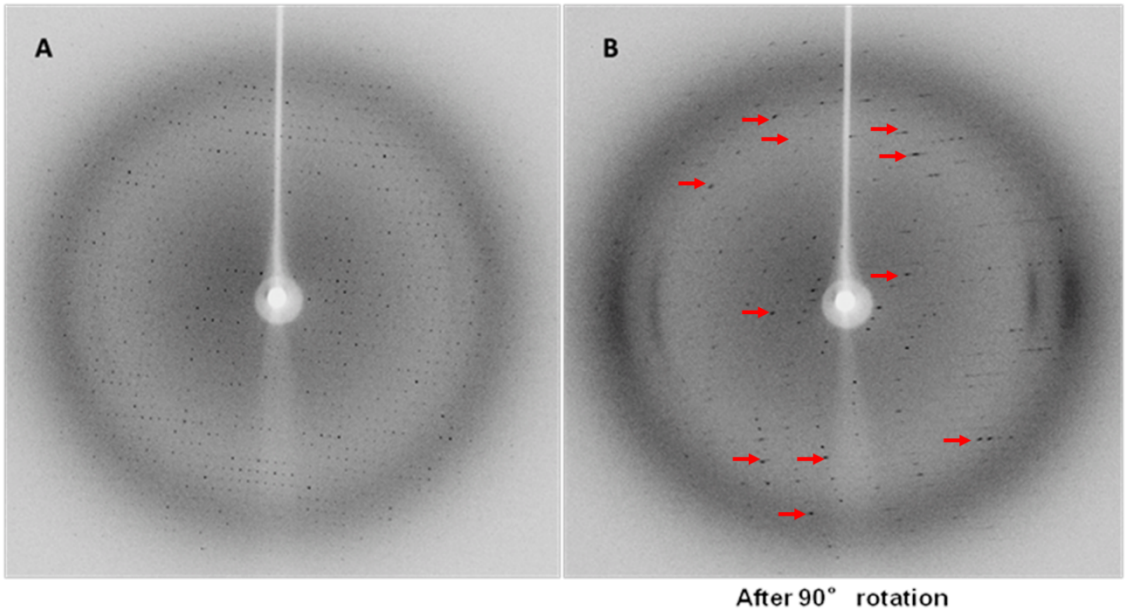


Figure 4.22: X-ray diffraction pattern reveals anisotropy. Red arrows show smeared reflections as an indication of anisotropy.

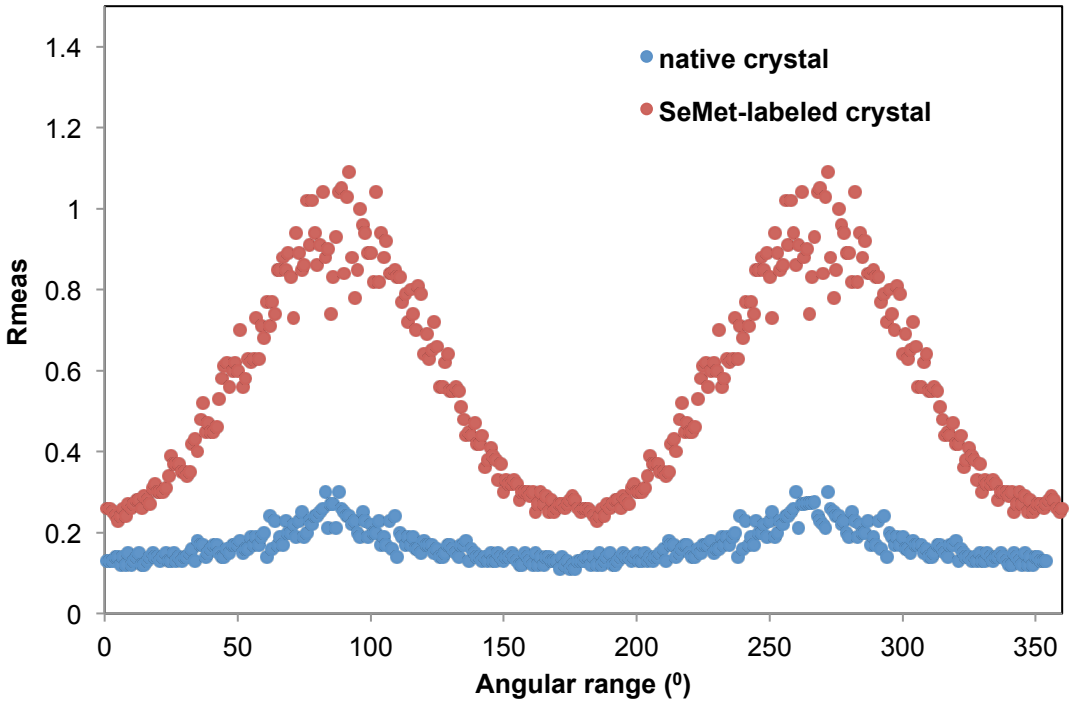


Figure 4.23: Rmeas analysis (Rupp 2010). It should be noticed that the native and derivative data sets were collected at two different beamlines. The native dataset was collected at micro-focus beam, while the derivative one was collected at normal beam-line.

4.9 Tackling anisotropy by changing external parameters

To solve anisotropy issue, external parameters were varied. Temperature, seeding, precipitation screen and pH screen were tried. Additive screens and silver bullet have been also tried in the attempt to obtain better diffraction quality crystals or crystals of different morphology. No significant improvement was obtained.

4.10 Co-crystallization of the protein with its substrates

Obtaining structure of PlaB in the complex with its substrate is one of the goals of this research project. Besides, ligand binding might stabilize the protein (Dekker, 1999) and results in better crystal packing, which may overcome twinning or anisotropy issues. Thus co-crystallization of PlaB D203N with its substrates was carried out. Point mutation D203N results in an inactive enzyme but may not alter the structure of the protein, due to similarity in structures of D and N. Therefore it can be assumed that the enzyme will not cleave the substrate, leaving the substrate intact in the binding pocket. It has not been

confirmed if the substrate really remains in the enzyme.

For co-crystallization of PlaB with its substrates, the substrates (PC, LPC, PG) were dissolved in either in Triton X-100, or 1.1% Octyl-β-D-glucopyranoside (OG), or 5% DMSO. Triton X-100 was used to dissolve the lipids for functional assay of PlaB (Kuhle, personal communication). OG is a nonionic detergent that is very common in crystallization of membrane protein and liposome reconstitution (Newby et al. 2009; Raman et al. 2006; Ferrandon & Newstead 2008). DMSO is a common solvent for dissolving organic molecules. The substrates were soluble in Triton X-100 and DMSO. 1.1% OG could not dissolve PC and PG completely. However, the best crystal hits were obtained when the substrates were dissolved in OG, although the size of the crystals was not sufficient for X-ray measurements and no new crystal morphology was found (Figure 4.24).

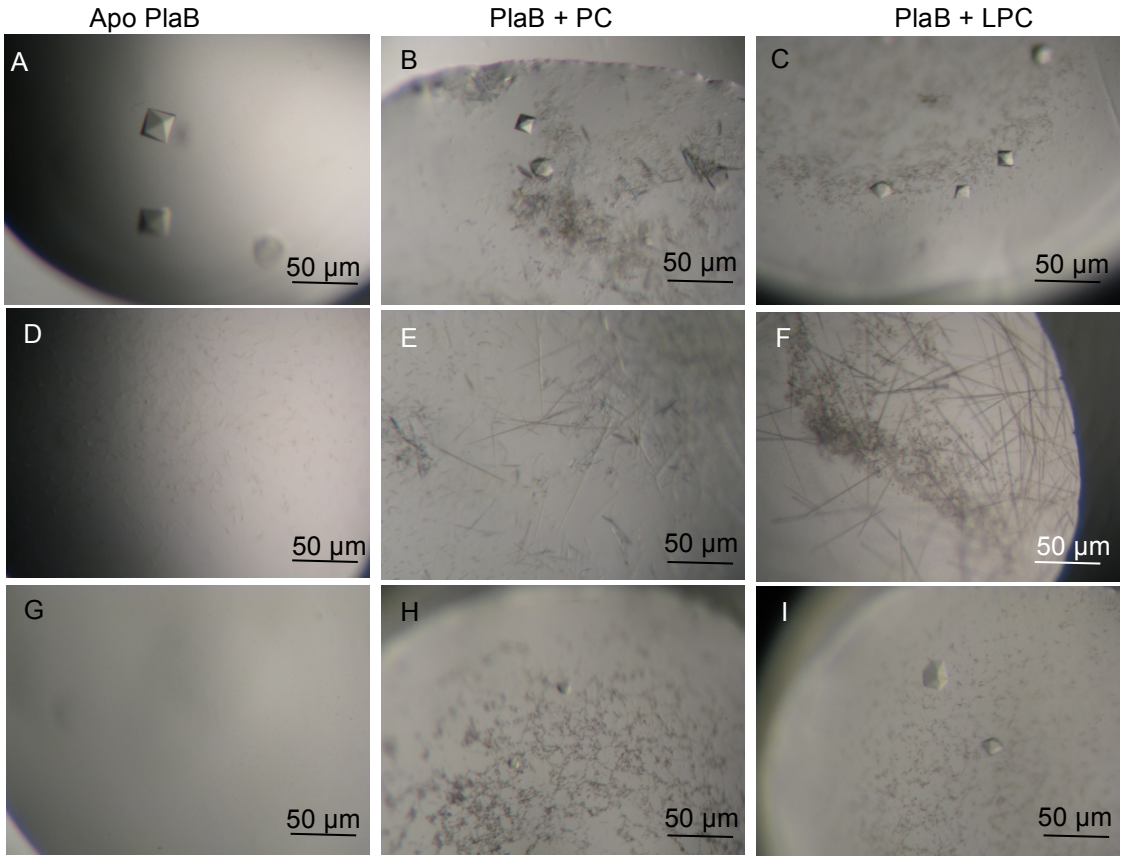


Figure 4.24: Initial hits of PlaB co-crystallized with its substrates dissolved in 1.1% OG. No crystals of PlaB in complex with PG. Crystallization conditions: For A, B, C: 0.1 M MES pH 6.5 and 15% (w/v) PEG 400; D, E, F: 0.2 M Ammonium sulfate, 0.1 M Bis-Tris pH 5.5 and 25% (w/v) PEG 3350. G, H, I: 0.2 M Ammonium acetate, 0.1 M Bis-Tris pH 5.5 and 25% PEG 3350.

4.11 Crystallization of chemically-modified protein

4.11.1 Lysine methylation

Various attempts focusing on external parameters to remove twinning and anisotropy issues have not resulted in positive outcomes. Chemical modifications of proteins were thus considered. The first, most simple approach is lysine methylation. Lysines are generally exposed on the surfaces of proteins and can disrupt well-ordered crystal lattice owing to its flexible amino acid side chains. Chemical modification of surface lysine residue help reduce surface entropy and thus may improve crystal packing, resulting in better diffracting crystals (Kobayashi et al. 1999; Kurinov et al. 2000; Rayment et al. 1993; Rayment 1997). The most common approach is the reductive methylation of free amino groups in which primary amines (i.e., lysine residues and the N terminus) are modified to tertiary amines.

The reaction was carried out according to the protocol (Walter et al. 2006). The methylated product was then purified by SEC (Figure 4.25A). 25% of protein amount remained soluble after lysine methylation (Figure 4.25B). The protein was concentrated down and initial crystallization screen was performed. However, crystallization of lysine-methylated protein only resulted in shower of needles (Figure 4.26). A grid screen of pH/PEG concentration did not improve the results.

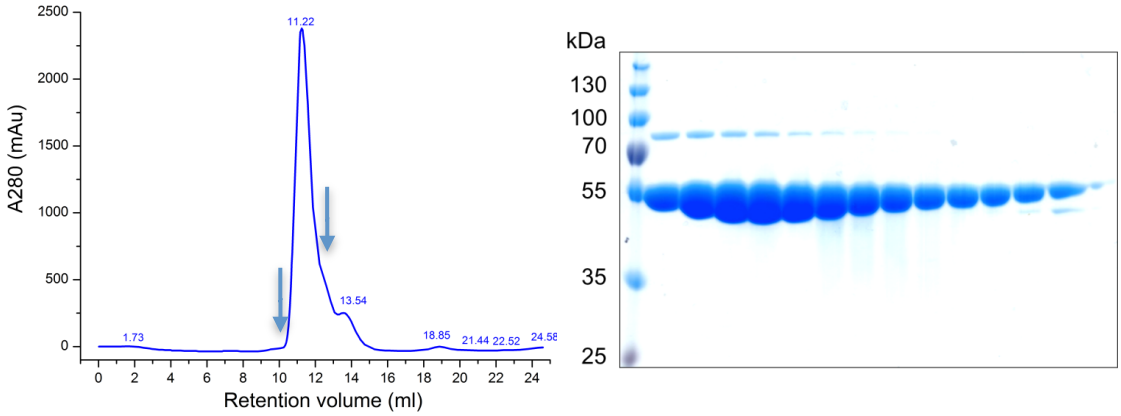


Figure 4.25: Purification of methylated lysine protein. A: Chromatogram of SEC run. Absorbance at 280nm is shown in blue. The arrows indicate the area that the elution fractions were collected. B. SDS-PAGE analysis of fractions collected from SEC purification of the protein after lysine methylation. Yield of the protein after lysine methylation reduced from 6 mg to 1.44 mg (25% remained after lysine methylation)

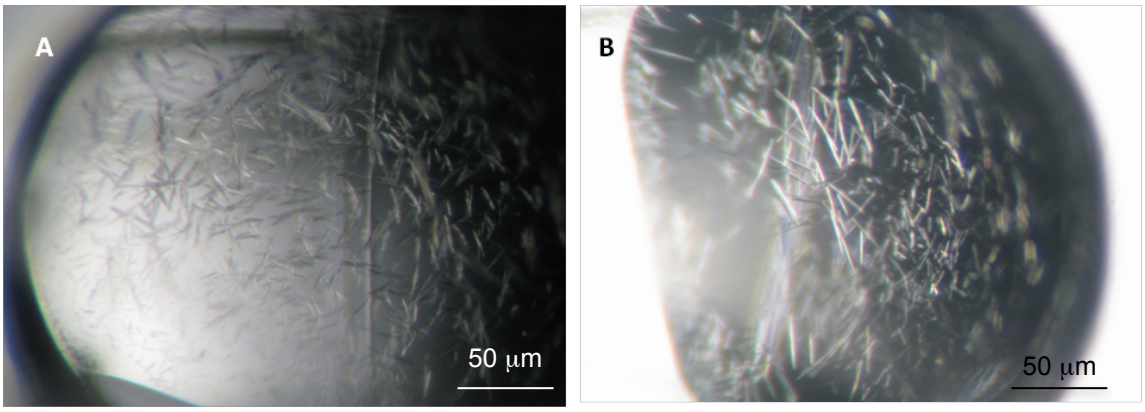


Figure 4.26: Lysine-methylated crystals. Crystallization conditions: A: 0.1 M MES pH 6.5, 15% (w/v) PEG 400; B: 2.4 M Sodium malonate pH 7.0.

4.11.2 Expression, purification and crystallization of C-terminal truncated protein

As mentioned above, the aim of the thesis is to determine the structure of a biological active enzyme (Although the working full length protein is the inactive mutant, it is assumed that there would be no significant difference in structure between wild type and D203N owing to similar structure of the two amino acids D and N). Truncation of only 15 amino acids at the C terminus of PlaB is sufficient to abolish the enzyme activity. Structure of a truncated enzyme thus may not give the answer on the enzyme's activity mechanism.

However, consensus secondary structure prediction at NPS@ (Deléage et al. 1997) suggests that the last 15 amino acids may be random coil (Figure 4.27). Removing this part may reduce flexibility of the protein which improves crystal packing. If a structure of a truncated version of PlaB in which the last 15 amino acids are deleted can be obtained, it can serve as a search model for molecular replacement to solve the structure of the full-length protein.

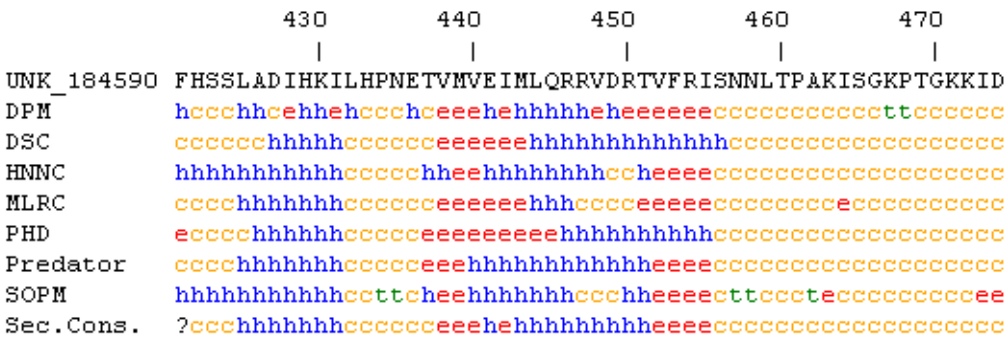


Figure 4.27: Consensus secondary structure prediction of PlaB's C-terminus. UNK_184590 is a designated id of PlaB sequence. DPM, DSC, HNNC, NLRC, PHD, Predator, SOPM are programs for secondary structure prediction. Sec.Cons. stands for secondary structure consensus.

Thus a truncated version of PlaB in which the last 15 amino acids were deleted (hereby designated as D203N_459) was expressed and purified following the protocol for full length protein (Figure 4.28). The protein was monodisperse on gel filtration with calculated molecular weight of 181.6 kDa derived from function given in Table 4.2. This allows estimation that the truncated protein is tetramer.

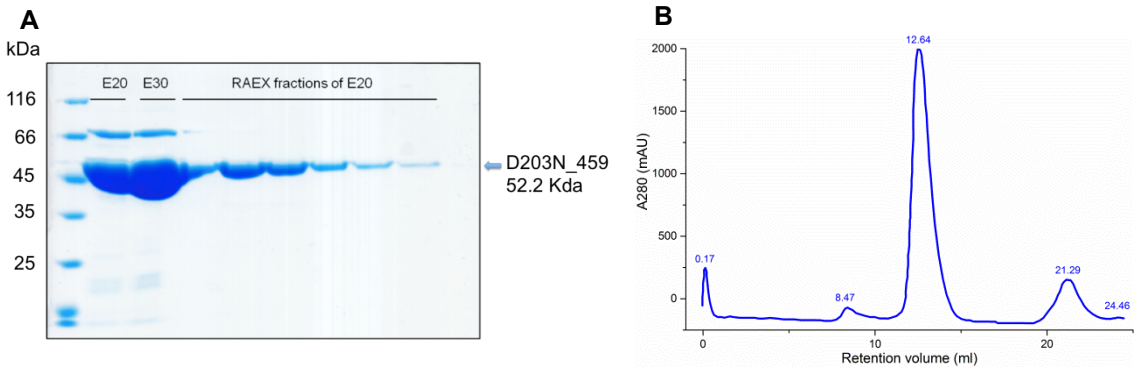


Figure 4.28: Analysis of C-terminal truncated PlaB after purification. A: SDS-PAGE analysis. E20 and E30 are the elution fractions after affinity purification. The number 20 and 30 mean the protein was expressed at 20°C and 30°C, respectively. Although the protein yield was higher when being produced 30°C, the protein is prone to aggregation. RAEX fractions are the protein fractions collected during reverse anion exchange chromatography. B: SEC analysis of RAEX fractions.

Purified protein was screened against various commercial crystallization kits. Crystals with different morphology were obtained (Figure 4.29), giving hope that the crystals may not have twinning or anisotropy problems. However the

crystals were still too small for x-ray measurements. Further optimization is needed to improve the size of the crystals.

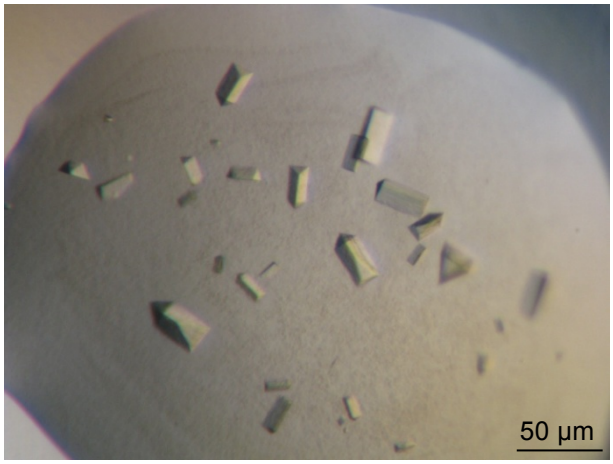


Figure 4.29: Initial crystals of D203N_459. The crystallization condition was 0.1 M HEPES pH 7.5, 20% PEG 6000. The full length protein also crystallized in either HEPES pH 7.5 or PEG 6000 and formed triclinal crystals. However, the crystal of D203N_459 has different morphology compared with full length protein.

4.11.3 Proteolysis screening

Limited proteolysis was used to identify globular and stable protein domains (Wernimont & Edwards 2009). The full length and C-terminus truncated proteins were treated with the proteases in Hampton kit as described in section 3.2.3.4. The products were analyzed by SDS-PAGE.

As shown in Figure 4.30, there is only a small portion of the protein was susceptible to different proteases, indicating by a smear pattern containing multi bands of lower MW with much weak intensity comparing with the full length band. The construct D203N_459 seems to be more stable. There was almost no degradation after 4 hours incubated with EG-C, EL, PA and PE (Figure 4.30B), while a small portion of the full length one was cleaved (Figure 4.30A). No stable sub-domain of PlaB was obtained for isolation and crystallization from these experiments.

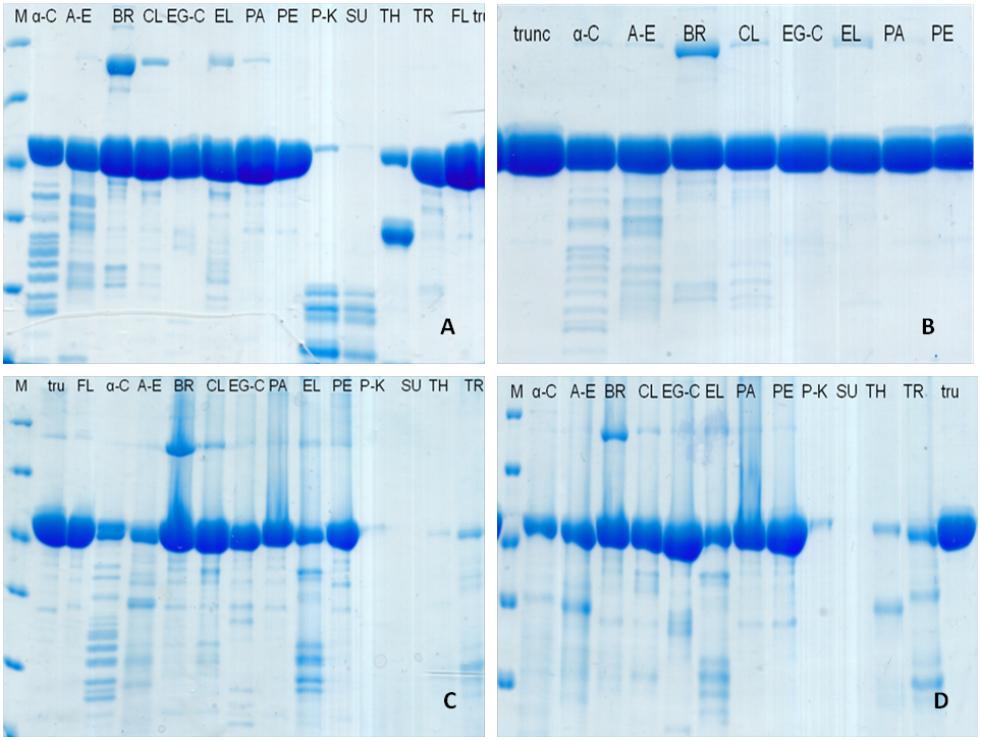


Figure 4.30: Proteolysis screening of full length (A & C) and C-terminus truncated proteins (B & D). The protein was incubated with enzymes for: (A) & (B): 4 hours; (C) & (D): overnight. FL: full length protein; tru: D203N_459.

4.11.4 Surface entropy reduction (SER) mutants

To identify SER mutants for construction, two approaches were tried. The first approach that was carried out by Kuhle relied on SERp Server (Goldschmidt et al. 2007). The second approach was to design all mutations around regions that are rich of hydrophilic amino acids (Cooper et al. 2007; Longenecker et al. 2001).

SER mutants suggested by SERp Server

These two surface mutants were constructed by K. Kuhle. In the first mutant, cluster EHKE (306 - 309) were changed to SHSS. In the second mutant, cluster KQGK (29 – 32) was mutated to SSGS. The proteins were expressed and purified according to the protocol for wild type PlaB. These two mutants, unfortunately, were heavily aggregated after affinity purification (Figure 4.31). The final concentration of the protein was 0.7 mg/ml which is usually insufficient for protein crystallization. An initial crystallization screen was still carried out but did not produce any crystals.

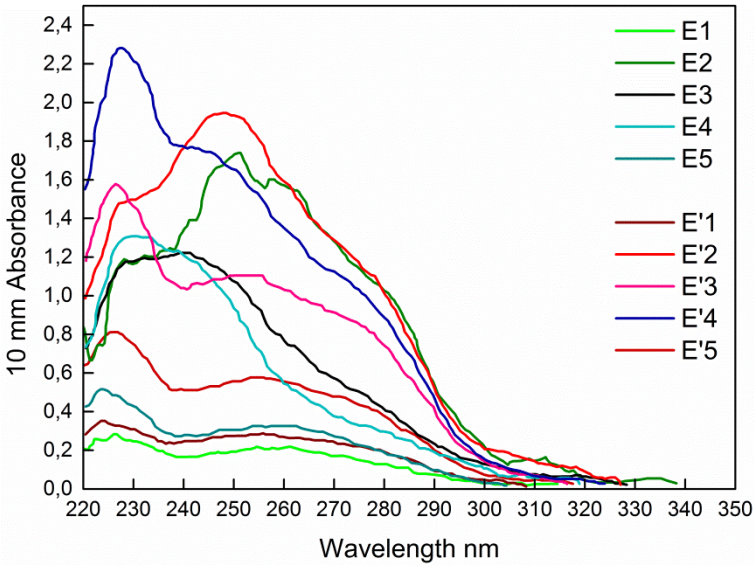


Figure 4.31: Absorbance scan of elution fractions of two surface mutants purified from Strep-tag affinity column. There were no peaks seen at A 280 nm in all elution fractions, suggesting that the proteins were precipitated.

Proposal for rational mutagenesis of predicted surface residues

Hydrophilic residues that usually reside on the surface of a protein were chosen for mutagenesis (Figure 4.32). 20 mutants were decided (Table 4.6) and primers for these 20 mutants were synthesized. Generation of these mutants is under way.

MIVIFVHGWSVTHTNTYG**E**LPQWL**E**N**QSKQGK**LDIQVGNIYLGRYISFDDT
VTVDDIARAFDQAVRD**E**IADKLRDGQRFACITHSTGGPIVRKWMDLYFKSN
LAKCPLSHLIMLAPANHGSALAQLGKSRLGRIKSFF**E**GI**E**PG**K**CVLDWL**E**LG
SDMSWQLN**E**SWLDYDCTANGVYSFVLTGQKIDRQFYDAVNSYTG**E**SGSD
GVVRVAATNMNYSLLKLHQ**E**GDNG**E**SLVVAKMTRTQPMAFGVLPGLSHS
GKNIGIIRSITMANAATHPTAIWILRCLQVKSRDSYNKLVK**E**LDNIT**KETQKN**
EHKEFVKTLVFTR**E**YITNRYSMIIFRLIDDRGNHLIDYDLYLTAGPQYS**EQ**AL
PAGFFVDRQRNLNNRGKLTYFLDYDIM**E**GGINTPKMQGNLGF**R**V**K**AYP**E**S
SDQALAYYRLLDFHSSLADIHKILHPN**E**TVMV**E**IMLQRRVDRTVFRISNNLT
PAKISGKPTG**KK**ID

Figure 4.32: Sequence of full length PlaB. Bold letters are residues predicted to be on the surface of the protein.

Table 4.6: Proposed SER mutants

K29AQ30AK32A	K29SQ30SK32S
E142AK144A	E142SK144S
K292AE293A	K292SE293S
K299AE300A	K299SE300S
E300AQ302AK303A	E300SQ302SK303S
K303AE305A	K303SE305S
K307AE308A	K307SE308S
E355AQ356A	E355SQ356S
R401AK403A	R401SK403S
K471AK472A	K471SK472S

5 Discussion

PlaB is a member of a novel lipase family (Bender et al. 2009) and is the most prominent cell-associated phospholipase A of *L.pneumophila* (Schunder et al. 2010). It exhibits both PLA and LPLA activity, and has hemolysis potential, through which it exposes its virulence. The enzyme also possesses unique features in the amino acid environment of catalytic residues which are not found in any characterised lipases. Because no homolog structure is available, structural determination of PlaB would set a foundation to visualize the activity of the protein and furthermore, to understand an important step in pathogenesis of *L. pneumophila*.

For this purpose, X-ray crystallography was chosen. Because truncation of only 15 amino acids at the C-terminus abolishes enzyme activity, attempts were spent on full-length protein to get structure of a biologically active protein. A preparation scheme for full length PlaB was established, the protein was characterised, native data set of full length protein was obtained, and various attempts were spent to obtain a derivative crystal that is qualified for phase solution step.

5.1 Purification and characterization of PlaB

The purification and characterization of the protein is the first and also the single most important step in working flow for crystallization (McPherson 1999).

Structure of a crystal is formed by naturally chemical bonding between molecules. Crystallization thus requires that the sample must be pure and homogeneous to ensure no interference by impurities or unnatural forms of the protein molecules (eg. aggregation). A study showed that proteins of unknown structure with purity level of > 95% judged by SDS-PAGE were successfully crystallized in 59% of all cases, whereas samples that were < 95% pure yield crystals with a success rate of only 37% (Geerlof et al. 2006). However, there is still no certain rule of how pure is pure enough. In the case of eubacterial aspartyl-tRNA synthetase, a purity level of ~99% is required for the production of high-quality crystals (Moreno et al. 2005). An amount of less than 5% of AcrB, which was not detectable on SDS-PAGE, as a contaminant in preparation of another target protein was sufficient for the protein to crystallize, while the target protein failed to crystallize (Veesler et al. 2008).

A preparation scheme for PlaB was established yielding sufficient amounts of protein qualified for crystallization. Factors that can affect the solubility of the protein were investigated. It is not uncommon for one batch of protein crystallizes while the next will not, and for many cases, the factor(s) that decides this difference can only be identified once the protein structure is solved (Mc Ree 1999), or takes a remarkable time to find out. It happened in the case of PlaB, which showed different behaviors among protein batches. The most effective preparation scheme not only must ensure a reasonably high amount of sample for crystallization at one time, but also should be as simple as possible.

5.2 PlaB is not degraded but tends to aggregate at higher temperature

SDS-PAGE analysis revealed that the full length PlaB was not degraded at R.T even after two weeks, either in solution or in crystal. However, DLS measurement was impossible for samples at R.T because the protein formed aggregation that is out of range for DLS measurement. DLS measurement could only be obtained when sample was at 4°C. Crystallization at 4°C resulted in no crystals but high numbers of clear drops, suggesting the protein was unable to aggregate in the time frame that was sufficient for protein crystallized at 20°C.

There has been no biological relevance with this trend found in previous studies on phospholipases. The tendency for aggregation of PlaB may not be a physiological process but purely due to *in vitro* experimental conditions. High temperature increases the energy of motion of the protein chain, which may break relatively weak H-bonds, as well as electrostatic and hydrophobic interactions (Lebendiker & Danieli 2014), resulting in the collapse of the protein conformation and thus protein became aggregated.

5.3 Oligomerization of PlaB in biological context

Gel filtration of PlaB at high concentration suggested that in *in vitro* solution, the protein formed a stable tetramer. This observation is supported by analytical ultra-centrifuge analysis (Kuhle, personal communication). At lower concentration, the protein dissociated to dimer, suggesting that oligomerization

of PlaB is concentration-dependent.

In nature, there are various examples in which the oligomerization interplay of a phospholipase controls its activity. For example, *Staphylococcus aureus* phosphatidylinositol-specific phospholipase C is inactive at monomer form, whereas dimerization leads to its activation. In the inactive state, soluble anions or anionic lipids bind to anion binding pockets of the enzyme and inhibits enzyme dimerization. PC activates the enzyme by binding to the lipase, therefore diluting anion substrate and allowing transient dimers to form (Cheng *et al.* 2012). *E.coli* outer membrane phospholipase A (OMPLA) present as an inactive monomer in the outer membrane of the bacteria. Phospholipid substrates activate the enzyme activity by inducing dimerization of the enzyme, resulting in the complete formation of substrate binding pockets and functional oxyanion holes (Stanley *et al.* 2007; Stanley *et al.* 2006). Usually, oligomerization is the activation mechanism of phospholipases.

However, PlaB does not follow the above mechanism. Although the tetramer form of the protein is stable and monodisperse in gel-filtration, indicating that this is not a non-specific interaction, the protein is most active in monomer form indicated by in-vitro assay (Kuhle *et al.* 2014).

PlaB does not participate in the internalization of *Legionella pneumophila* but rather at the bacterial replication and dissemination steps. So there must be a mechanism that regulates its activity and keeps the enzyme inactive at certain point in bacterial life cycle. It may be hypothesized that the tetramer form is the storage state of the enzyme. The binding of the substrates may lead to conformational change of the oligomer. The tetramer disassociate to monomer, resulting in the exposure of the active site. Further experiments are needed to confirm this hypothesis.

Independently, by using biophysical and enzymatic essays, Kuhle *et al.* had also come to similar suggestion that oligomer may inhibits PlaB's activity (Kuhle *et al.* 2014).

As indicated in section 4.11.2, the protein was estimated to be tetramer after the truncation of the last 15 amino acids at C-terminus. This suggests that the last 15 amino acids, although play an on/off role in enzyme activity, do not contribute to the oligomerization of the enzyme.

5.4 Issues in crystallization

Obtaining high quality crystals is a critical step for but has always been a bottle neck to protein structure determination by X-ray crystallography. To obtain good structural data, crystals need to be single and isotropic. Although it was able to obtain the PlaB protein at high yield and high homogeneity after several optimizations, which is supposed to be suitable for crystallization, crystals were obtained with reasonable resolutions, the crystals has always been perfectly twinned in hexagonal form or anisotropic in triclinic form. There is no clear rule on which condition gives hexagonal crystals, or which condition gives triclinic crystals. Both low MW and high MW PEG can give either hexagonal or triclinic crystals. Although hexagonal crystals did not appeared at pH above 8, rod-shaped crystals appeared in a wide range of pH from 6.5 to above 8. It was even observed sometimes that a mix of hexagonal and rod-shaped crystals appeared in the same condition that had given hexagonal ones before. This behavior creates difficulties for optimization focus. For example, one may want to focus on solving twinning issue of the hexagonal crystals, but when repeating the crystallization condition and expanding optimization around the condition, the crystals turned out to be triclinic one.

Other issues in crystallization of PlaB include quick aggregation that made crystal form just after 5 minutes, irreproducibility of crystallization behaviors between different batches of purified proteins, or even between different plates, pipetting of the same protein batch and crystallization condition.

6 Outlook

Vapor diffusion techniques were employed in this work to grow the protein crystals. In these techniques, evaporating the solvent from a droplet containing the purified protein buffered at a given pH and precipitant is the key to reach the supersaturation state needed for crystallization to occur. The evaporation of the solvent from the droplet occurs during equilibration of the droplet with a larger reservoir containing a solution at a higher precipitant concentration. The equilibrium state is attained when the vapor pressure at the droplet surface balances with the vapor pressure of the reservoir. At this point of time, the supersaturation is no longer affected by the solvent evaporation.

Obtaining well diffracting crystals is in a great deal determined by solvent evaporation and thus by parameters governing the evaporation process. These parameters include the temperature, drop size and shape, and concentrations of crystallizing agent in the droplet and dehydrating agent in the reservoir (Boistelle & Astier, 1988; Fowlis et al., 1988; Mikol et al., 1990; Sibille et al., 1991).

Different approaches were tested in the attempt to control the rate of solvent evaporation. These include testing different concentrations and different types of PEGs, different pH, different crystallization temperatures, additive screen and silver bullets. These approaches did not produce a significant improvement.

Among parameters that affect solvent evaporation, drop shape was not taken into account, and this maybe an important cause of protein crystallization behavior. This parameter, however, is the most difficult factor to reproduce. It heavily depends on manual handling, the material of the plates and the composition of the crystallization contents. While the material of the plates and composition of the crystallization contents can be keep constant during crystallization experiments, it is a challenge to make manual handling reproducible.

Among chemical-modification approaches that have been tried to intervene the chemical nature of the protein and eventually to change crystal packing, a different crystal morphology was obtained by crystallizing a truncated version of PlaB where the last 15 amino acids, predicted to be disordered, were cut off. A different morphology gives hope that the crystal may not have twinning or anisotropy issue. This condition should be optimized further to obtain a crystal

with diffracting quality. If a structure of this version can be solved, it can be used as search model for determining structure of full-length protein whose native data set is available.

Future work should focus on optimizing the crystals of the C-terminal truncated PlaB and crystallization of SER mutants to overcome twinning and anisotropy issues. It would then be hopeful that PlaB structure could be solved and establish a knowledge foundation for new drug discovery and to fully understand molecular mechanism of PlaB's activities.

7 References

Agarwal, A.K. & Fishwick, C.W.G. (2010). Structure-based design of anti-infectives. *Annals of the New York Academy of Sciences*, 1213(1), 20–45.

Albert-Weissenberger, C., Cazalet, C. & Buchrieser, C. (2007). Legionella pneumophila - a human pathogen that co-evolved with fresh water protozoa. *Cellular and molecular life sciences: CMLS*, 64(4), 432–48.

Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics (Oxford, England)*, 22(2), 195–201.

Arpigny, J.L. & Jaeger, K.E. (1999). Bacterial lipolytic enzymes: classification and properties. *The Biochemical journal*, 343 Pt 1, 177–83.

Banerji, S., Aurass, P. & Flieger, A. (2008). The manifold phospholipases A of *Legionella pneumophila* - identification, export, regulation, and their link to bacterial virulence. *International journal of medical microbiology: IJMM*, 298(3-4), 169–81.

Barker, J., Scaife, H. & Brown, M.R.W. (1995). Intraphagocytic growth induces an antibiotic-resistant phenotype of *Legionella pneumophila*. Intraphagocytic Growth Induces an Antibiotic-Resistant Phenotype of *Legionella pneumophila. Antimicrobial agents and chemotherapy*, 39(12), 2684-88.

Beddell, C.R., Goodford P.J., Norrington, F.E. & Wootton, R. (1976). Compounds designed to fit a site of known structure in human haemoglobin. *British journal of pharmacology*, 57(2), 201–9.

Bender, J., Rydzewski, K., Broich, M., Schunder, E. & Heuner, K. (2009). Phospholipase PlaB of *Legionella pneumophila* represents a novel lipase family: protein residues essential for lipolytic activity, substrate specificity, and hemolysis. *The Journal of biological chemistry*, 284(40), 27185–94.

Bernhard Rupp (2010). Biomolecular Crystallography: Principles, practicem and Application to Structural Biology. Garland Science, Taylor & Francis Group, LIC.

Blow, D. & Rossmann, M. (1961). The single isomorphous replacement method. *Acta crystallographica*, 14(11), 1195-1202.

Boggon, T.J. & Shapiro, L. (2000). Screening for phasing atoms in protein crystallography. *Structure (London, England: 1993)*, 8(7), R143–9.

Boistelle, R. & Astier, J.P. (1988). Crystallization mechanisms in solution. *Journal of Crystal Growth*, 90(1-3), 14–30.

Borgstahl, G.E.O. (2007). How to use dynamic light scattering to improve the likelihood of growing macromolecular crystals. *Methods in molecular biology (Clifton, N.J.)*, 363, 109–29.

De Buck, E., Anné, J. & Lammertyn, E. (2007). The role of protein secretion systems in the virulence of the intracellular pathogen *Legionella pneumophila*. *Microbiology (Reading, England)*, 153 (Pt 12), 3948–53.

Dodson, E. (2008). The befores and afters of molecular replacement. Acta crystallographica D, 64(1), 17-24.

Carratalà, J; Garcia-Vidal, C. (2010). An update on *Legionella*. *Current Opinion in Infectious Diseases*, 23, 152–157.

Chayen, N.E. (2004). Turning protein crystallization from an art into a science. *Current opinion in structural biology*, 14(5), 577–83.

Cheng, A.C., Coleman, R.G., Smyth, K.T., Cao, Q., Soulard, P., Caffrey, D.R., Salzberg, A.C. & Huang, E.S. (2007). Structure-based maximal affinity model predicts small-molecule druggability. *Nature biotechnology*, 25(1), 71–5.

Cheng, J., Goldstein, R., Stec, B., Gershenson, A. & Roberts, M.F. (2012). Competition between anion binding and dimerization modulates *Staphylococcus aureus* phosphatidylinositol-specific phospholipase C enzymatic activity. *The Journal of biological chemistry*, 287(48), 40317–27.

Chopra, I. (2013). The 2012 Garrod lecture: discovery of antibacterial drugs in the 21st century. *The Journal of antimicrobial chemotherapy*, 68(3), 496–505.

Cooper, D.R., Boczek, T., Grelewska, K., Pinkowska, M., Sikorska, M. & Derewenda. Z. (2007). Protein crystallization by surface entropy reduction: optimization of the SER strategy. *Acta crystallographica. Section D, Biological crystallography*, 63(Pt 5), 636–45.

Deléage, G., Blanchet, C. & Geourjon, C. (1997). Protein structure prediction. Implications for the biologist. *Biochimie*, 79(11), 681–6.

Diederen, B.M.W. (2008). Legionella spp. and Legionnaires' disease. *The Journal of infection*, 56(1), 1–12.

Diseases, M. (1968). Solvent Content of Protein. *Journal of molecular biology*, 33(2), 491–497.

Edelstein, P.H. & Meyer, R.D. (1980). Susceptibility of Legionella pneumophila to twenty antimicrobial agents. *Antimicrobial agents and chemotherapy*, 18(3), 403–8.

Erickson, J., Neidhart D.J., VanDrie J., Kempf D.J., Wang, X.C., Norbeck, D.W., Plattner, J.J., Rittenhouse J.W., Turon, M. & Wideburg, N. *et al*. (1990). Design, activity, and 2.8 A crystal structure of a C2 symmetric

inhibitor complexed to HIV-1 protease. *Science (New York, N.Y.)*, 249(4968), 527–33.

Evans, P. & McCoy, A. (2008). An introduction to molecular replacement. Acta crystallographica D 64(1): 1-10.

Ferhat, M., Atlan, D., Vianney, A., Lazzaroni, J-C., Doublet, P. & Gilbert, C. (2009). The TolC protein of Legionella pneumophila plays a major role in multi-drug resistance and the early steps of host invasion. *PloS one*, 4(11), e7732.

Ferrandon, B. & Newstead, S. (2008). Rationalizing alpha-helical membrane protein crystallization. *Protein Science*, 17, 466-72.

Fields, B.S. (1996). The molecular ecology of *legionellae*. *Trends in microbiology*, 4(7), 286–90.

Fields, B.S., Benson, R.F. & Besser, R.E. (2002). *Legionella* and Legionnaires' Disease: 25 Years of Investigation *Legionella* and Legionnaires' Disease: 25 Years of Investigation. Clinical Microbiology Reviews , 15(3), 506-26.

Flieger, A., Gong, S., Faigle, M., Deeg, M., Bartmann, P. & Neumeister, B. (2000). Novel phospholipase A activity secreted by *Legionella* species. *Journal of bacteriology*, 182(5), 1321–7.

Flieger, A, Gongab, S., Faigle, M., Mayer, H.A., Kehrer, U., Mussotter, J., Bartmann, P. & Neumeister, B. (2000). Phospholipase A secreted by *Legionella pneumophila* destroys alveolar surfactant phospholipids. *FEMS microbiology letters*, 188(2), pp.129–33.

Flieger, A., Rydzewski, K., Banerji, S., Broich, M. & Heuner, K. (2004). Cloning and Characterization of the Gene Encoding the Major Cell-Associated Phospholipase A of *Legionella pneumophila*, *plaB*, Exhibiting Hemolytic Activity. *Infection and immunity*, 72(5), 2648–58.

Fliermans, C.B., Cherry, W.B., Orrison, L.H., Smith, S.J., Tison, D.L. & Pope, D.H. (1981). Ecological distribution of Ecological Distribution of *Legionella pneumophila*. *Applied and Environmental Microbilogy*, 41(1), 9–16.

Fong, D.H., Lemke, C.T., Hwang, J., Xiong, B. & Berghuis, A.M. (2010). Structure of the antibiotic resistance factor spectinomycin phosphotransferase from *Legionella pneumophila*. *The Journal of Biological Chemistry*, 285(13), 9545–55.

Fowlis, W.W., DeLucas, L.J., Twigg, P.J., Howard, S.B., Meehan Jr., E.J. & Baird J.K. (1988). Experimental and theoretical analysis of the rate of solvent equilibration in the hanging drop method of protein crystal growth. *Journal of Crystal Growth*, 90(1-3), 117–129.

Fraser, D.W. (1985). Potable water as a source for legionellosis. *Environmental Health Perspectives*, 62, 337–41.

Fraser, D.W., Tsai, T.R., Orenstein, W., Parkin, W.E., Beecham, H.J., Sharrar, R.G., Harris, J., Mallison, G.F., Martin, S.M., McDade, J.E., Shepard, C.C., Brachman, P.S. (1997). Legionnaires' Disease - Description of an Epidemic of Pneumonia. *The New England Journal of Medicine*, 297, 1189–97.

Garman, E.F. (2010). Radiation damage in macromolecular crystallography: what is it and why should we care? *Acta crystallographica. Section D, Biological crystallography*, 66 (Pt 4), 339–51.

Geerlof, A., Brown, J., Coutard, B., Egloff, M.P., Enquita, F.J., Fogg, M.J., Gilbert, R.J., Groves, M.R., Haouz, A., Nettleship, J.E., Nordlund, P., Owens, R.J., Ruff, M., Saisbury, S., Svergun, D.I. & Wilmanns, M. (2006). The impact of protein characterization in structural proteomics. *Acta crystallographica. Section D, Biological Crystallography*, 62b(Pt 10), 1125–36.

Godoy, L.C., Duquesne, S., Bordes, F., Sandoval & G., Marty., A. (2012). Lipases: An Overview. *Methods in Molecular Biology*, 861, 3–30.

Goldschmidt, L., Cooper, D.R. & Derewenda, Z.S. (2007). Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. 1569–76.

Green, D.W., Ingram, V. M. & Perutz, M.F. (1954). The Structure of Haemoglobin. IV. Sign Determination by the Isomorphous Replacement Method. *Proceedings of the Royal Society of London. Series A*, 225(1162), 287-307.

Halgren, T.A. (2009). Identifying and characterizing binding sites and assessing druggability. *Journal of chemical information and modeling*, 49(2), 377–89.

Hartmann, B.G.R., Heinrich, P., Kollenda, M.C., Skrobranek, B., Tropschug, M. & Weiβ, W. (1985). Molecular Mechanism of Action of the Antibiotic Rifampicin. *Angewandte Chemie International Edition in English*, 24 (12), 1009–14.

Heath, C.H., Grove, D.I. & Looke, D.F. (1996). Delay in appropriate therapy of *Legionella pneumonia* associated with increased mortality. *European journal of clinical microbiology & infectious diseases: official publication of the European Society of Clinical Microbiology*, 15(4), 286–90.

Hilbi, H., Jarraud, S., Hartland, E. & Buchrieser, C. (2010). Update on Legionnaires' disease: pathogenesis, epidemiology, detection and control. *Molecular microbiology*, 76(1), 1–11.

Hooper, D.C. (2001). Mechanisms of action of antimicrobials: focus on fluoroquinolones. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*, 32 Suppl 1 (Suppl 1), S9–S15.

Horwitz, B.Y.M.A., Blood, H. & Cells, M. (1983). The Legionnaires' disease bacterium (*Legionella pneumophila*) inhibits phagosome-lysosome fusion in human monocytes. *The Journal of Experimental medicine*, 158(December), 2108–2126.

Horwitz, M. A (1984). Phagocytosis of the Legionnaires' disease bacterium (*Legionella pneumophila*) occurs by a novel mechanism: engulfment within a pseudopod coil. *Cell*, 36(1), 27–33.

Horwitz, M. A & Maxfield, F.R. (1984). *Legionella pneumophila* inhibits acidification of its phagosome in human monocytes. *The Journal of cell biology*, 99(6), 1936–43.

Istivan, T.S. & Coloe, P.J. 2006. Phospholipase A in Gram-negative bacteria and its role in pathogenesis. *Microbiology (Reading, England)*, 152(Pt 5), 1263–74.

Jäger, J., Marwitz, S., Tiefenau, J., Rasch, J., Shevchuk, O., Kugler, C., Goldmann, T. & Steinert, M. (2014). Human lung tissue explants reveal novel interactions during *Legionella pneumophila* infections. *Infection and Immunity*, 82(1), 275–85.

Kabsch, W. (2010). Xds. *Acta crystallographica. Section D, Biological Crystallography*, 66(Pt 2), 125–32.

Kantardjieff, K.A. & Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins , DNA , and protein – nucleic acid complex crystals. *Protein Science*, 12(9), 1865–71.

Kauffmann, I. & Schmidt-Dannert, C. (2001). Conversion of *Bacillus thermocatenulatus* lipase into an efficient phospholipase with increased activity towards long-chain fatty acyl substrates by directed evolution and rational design. *Protein engineering*, 14(11), 919–28.

Kelley, L. A & Sternberg, M.J.E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols*, 4(3), 363–71.

Kobayashi, M., Kubota, M. & Matsuura, Y. 1999. Crystallization and improvement of crystal quality for X-ray diffraction of maltooligosyl trehalose synthase by reductive methylation of lysine residues. *Acta Crystallographica Section D Biological Crystallography*, 55(4), 931–933.

Kotra, L.P., Haddad, J. & Mobashery, S. (2000). Aminoglycosides: Perspectives on Mechanisms of Action and Resistance and Strategies to Counter Resistance. *Antimicrobial Agents and Chemotherapy*, 44(12): 3249-56.

Kuhle, K. & Flieger, A. (2013). *Legionella* Phospholipases Implicated in Virulence. *Current Topics in Microbiology and Immunology*, 376, 175–209.

Kurinov, I. V, Mao, C., Irvin, J.D., Uckun & F.M. (2000). X-ray crystallographic analysis of pokeweed antiviral protein-II after reductive methylation of lysine residues. *Biochemical and Biophysical Research Communications*, 275(2), pp.549–52.

Kwaik, Y.A., Gao, L.Y., Stone, B.J., Venkataraman, C. & Harb, O.S. (1998). Invasion of Protozoa by *Legionella pneumophila* and Its Role in Bacterial Ecology and Pathogenesis. *Applied and Environmental Microbiology*, 64(9), 3127-33.

Laemmli, U.K. (1970). Cleavage of Structural Proteins during the Assembly of the head of bacteriophage T4. *Nature*, 227, 680-5.

Lang, C. & Flieger, A. (2011). Characterisation of *Legionella pneumophila* phospholipases and their impact on host cells. *European Journal of Cell Biology*, 90(11), 903–12.

Lebendiker, M. & Danieli, T. (2014). Production of prone-to-aggregate proteins. *FEBS letters*, 588(2), 236–46.

Longenecker, K.L., Garrard S.M., Sheffield, P.J. & Derewenda, Z.S. (2001). Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of Rho GDI. *Acta Crystallographica Section D Biological Crystallography*, 57(5), 679–688.

Marston, B.J., Plouffe J.F., File T.M.Jr, Hackman B.A., Salstrom, S.J., Lipman H.B., Kolczak, M.S., Breiman, R.F. (1997). Incidence of community-acquired pneumonia requiring hospitalization: Results of a population-based active surveillance study in Ohio. The community-based pneumonia incidence study group. *Archives of Internal Medicine*, 157(15), 1709–1718.

Mc Ree, D.E. (1999). *Practical Protein Crystallography* 2nd ed., Academic Press.

McDade, J., Shepard, C., Fraser, D., Tsai, T., Redus, M. & Dowdle, W. (1977). Legionnaires' Disease — Isolation of a Bacterium and Demonstration of Its Role in Other Respiratory Disease. *The New England Journal of Medicine*, 297, 1197–1203.

McPherson, A. (1999). *Crystallization of Biological Macromolecules*, Cold Spring Harbor Laboratory Press.

Mikol, V., Rodeau, J.-L. & Giegé, R. (1990). Experimental determination of water equilibration rates in the hanging drop method of protein crystallization. *Analytical Biochemistry*, 186(2), 332–39.

Nguyen, T.M., Ilef, D., Jarraud, S., Rouil, L., Campesse, C., Che, D., Haeghebaert, S., Ganiayre, F., Etienne, J & Desenclos, J.C. (2006). A Community-Wide Outbreak of Legionnaires Disease Linked to Industrial

Cooling Towers — How Far Can Contaminated Aerosols Spread? *The Journal of Infectious Diseases*, 193(1), 102–11.

Moffie, B.G. & Mouton, R.P. (1988). Sensitivity and resistance of *Legionella pneumophila* to some antibiotics and combinations of antibiotics. *The Journal of antimicrobial chemotherapy*, 22(4), 457–62.

Molmeret, M., Bitar, D.M., Han, L. & Kwaik, Y.A. (2004). Cell biology of the intracellular infection by *Legionella pneumophila*. *Microbes and Infection*, 6(1), 129–39.

Molmeret, M. et al., 2007. Rapid escape of the dot/icm mutants of Legionella pneumophila into the cytosol of mammalian and protozoan cells. *Infection and immunity*, 75(7), pp.3290–304.

Molmeret, M., Santic, M., Asare, R., Carabeo, R.A. & Abu Kwaik, Y. (2010). Temporal and spatial trigger of post-exponential virulence-associated regulatory cascades by *Legionella pneumophila* after bacterial escape into the host cell cytosol. *Environmental microbiology*, 12(3), 704–15.

Moreno, A., Théobald-Dietrich, A., Lorber, B., Sauter, C. & Giegé, R. (2005). Effects of macromolecular impurities and of crystallization method on the quality of eubacterial aspartyl-tRNA synthetase crystals. *Acta crystallographica. Section D, Biological crystallography*, 61(Pt 6), 789–92.

Newby, Z.E., O'Connell J.D. 3rd, Gruswitz, F., Hays, F.A., Harries, W.E, Harwood, I.M., Ho, J.D., Lee, J.K., Savage, D.F., Miercke, L.J., Stroud, R.M. (2009). A general protocol for the crystallization of membrane proteins for X-ray structural investigation. *Nature protocols*, 4(5), 619–37.

Newton, H.J., Ang, D.K., van Driel, I.R. & Hartland, E.L. (2010). Molecular pathogenesis of infections caused by *Legionella pneumophila*. *Clinical microbiology reviews*, 23(2), 274–98.

Nicolini, A., Ferraioli, G. & Senarega, R. (2013). Severe *Legionella pneumophila* pneumonia and non-invasive ventilation: presentation of two cases and brief review of the literature. *Pneumonologia i alergologia polska*, 81(4), 399–403.

Osterholm, M.T, Chin, T.D., Osborne, D.O, Dull, H.B., Dean, A.G., Fraser, D.W., Hayes, P.S., Hall, W.N. (1983). A 1957 outbrake of Legionnaires' disease associated with a meat packing plant. *American Journal of Epidemiology*, 117(1), 60–67.

Padilla, J.E. & Yeates, T.O. (2003). A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. *Acta Crystallographica Section D Biological Crystallography*, 59(7), 1124–1130.

Proteau, A., Shi, R. & Cygler, M. (2010). Application of dynamic light scattering in protein crystallization. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]*, Chapter 17(August), p.Unit 17.10.

Raman, P., Cherezov, V. & Caffrey, M. (2006). The Membrane Protein Data Bank. *Cellular and molecular life sciences: CMLS*, 63(1), 36–51.

Rayment, I. (1997). Reductive alkylation of lysine residues to alter crystallization properties of proteins. *Methods in enzymology*, 276, 171–9.

Rayment, I., Rypniewski, W.R., Schmidt-Bäse, K., Smith, R., Tomchick, D.R., Benning, M.M., Winkelmann, D.A., Wesenberg, G., Holden, H.M. (1993). Three-Dimensional Structure of Myosin Subfragment- 1: A Molecular Motor. *Science*, 261, 50–58.

Roberts, N. A., Martin J.A., Kinchington, D., Broadhurst, A.V., Craig, J.C., Duncan, I.B., Galpin, S.A., Handa, B.K., Kay, J., Kröhn A, et al. (1990). Rational design of peptide-based HIV proteinase inhibitors. *Science (New York, N.Y.)*, 248(4953), 358–61.

Rossmann, M.G. (2001). Molecular replacement - historical background. Acta Crystallographica D, 57(10), 1360-1366.

Rossmann, M.G. & Blow, D.M. (1962). The detection of sub-units within the crystallographic asymmetric unit. Acta Crystallographica D, 15(1), 24-31.

Rowbotham, T.J., 1980. Preliminary report on the pathogenicity of *Legionella pneumophila* for freshwater and soil amoebae. *Journal of Clinical Pathology*, 33(12), 1179–83.

Muder, R.R., Yu, V.L., Woo, A.H. (1986). Mode of transmission of *Legionella pneumophila*: A critical review. *Archives of Internal Medicine*, 146(8), 1607–1612.

Sambrook, J & Russell, D.W. (2001). Molecular cloning: a laboratory manual. 3rd. ed. Cold Spring Habor Laboratory.

Schmidt, T.G.M. & Skerra, A. (2007). The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nature protocols*, 2(6), 1528–35.

Schmiel, D.H. & Miller, V.L. (1999). Bacterial phospholipases and pathogenesis. *Microbes and infection / Institut Pasteur*, 1(13), 1103–12.

Schunder, E. Adam, P., Higa, F., Remer, K.A., Lorenz, U., Bender, J., Schulz, T., Flieger, A., Steinert, M. & Heuner, K. (2010). Phospholipase PlaB is a new virulence factor of *Legionella pneumophila*. *International journal of medical microbiology: IJMM*, 300(5), 313–23.

Shi, J., Blundell, T.L. & Mizuguchi, K. (2001). FUGUE: sequence-structure homology recognition using environment-specific substitution tables and

structure-dependent gap penalties. *Journal of molecular biology*, 310(1), 243–57.

Sibille, L., Clunie, J.C. & Baird, J.K. (1991). Solvent evaporation rates in the closed capillary vapor diffusion method of protein crystal growth. *Journal of Crystal Growth*, 110(1-2), 80–88.

Simmons, K.J., Chopra, I. & Fishwick, C.W.G. (2010). Structure-based discovery of antibacterial drugs. *Nat Rev Micro*, 8(7), 501–510.

Stanley, A.M., Chuawong, P., Hendrickson, T.L., Fleming, K.G. (2006). Energetics of outer membrane phospholipase A (OMPLA) dimerization. *Journal of Molecular Biology*, 358(1), 120–31.

Stanley, A.M., Treubrodt, A.M., Chuawong, P., Hendrickson, T.L. & Fleming, K.G. (2007). Lipid chain selectivity by outer membrane phospholipase A. *Journal of Molecular Biology*, 366(2), 461–8.

Sugahara, M., Asada, Y., Shimada, H., Taka, H. & Kunishima, N. (2009). HATODAS II – heavy-atom database system with potentiality scoring. *Journal of Applied Crystallography*, 42(3), 540–4.

Thornsberry, C., Baker, C.N. & Kirven, L. A (1978). In vitro activity of antimicrobial agents on Legionnaires disease bacterium. *Antimicrobial Agents and Chemotherapy*, 13(1), 78–80.

Tyndall, J., Sinchaikul, S., Fothergill-Gilmore, L.A., Taylor, P. & Walkinshaw, M.D. (2002). Crystal Structure of a Thermostable Lipase from *Bacillus stearothermophilus* P1. Journal of Molecular Biology, 323(5), 859-869.

Vacca, J.P., Dorsey, B.D., Schleif, W.A., Levin R.B., McDaniel, S.L., Darke P.L., Zugay, J., Quintero, J.C., Blahy, O.M. & Roth, E. (1994). L-735,524: The Design of a Potent and Orally Bioavailable HIV Protease Inhibitor. *Journal of Medicinal Chemistry*, 37(21), 3443-51.

Van Kampen, M.D., Simons, J.W., Dekker, N., Egmond, M.R. & Verheij, H.M. (1998). The phospholipase activity of *Staphylococcus hyicus* lipase strongly depends on a single Ser to Val mutation. *Chemistry and Physics of Lipids*, 93(1-2), 39–45.

Van Pham S.T., Engman, H., Dahlgren, L.G., Cornvik, T., Eshaghi, S. (2010). A systematic approach to isolate mono-disperse membrane proteins - purification of zinc transporter ZntB. *Protein Expression and Purification*, 72(1), 48–54.

Veesler, D., Blangy, S., Cambillau, C. & Sciara, G. (2008). There is a baby in the bath water: AcrB contamination is a major problem in membrane-protein crystallization. *Acta crystallographica. Section F, Structural biology and crystallization communications*, 64(Pt 10), 880–5.

Villoutreix, B.O. Renault, N., Lagorce, D., Sperandio, O., Montes, M., Miteva M.A. (2007). Free resources to assist structure-based virtual ligand screening experiments. *Current Protein & Peptide Science*, 8(4), 381–411.

Walsh, C.T. (1983). Suicide substrates: mechanism-based enzyme inactivators with therapeutic potential. *Trends in Biochemical Sciences*, 8(7), 254–257.

Walter, T.S., Meier, C., Assenberg, R., Au, K.F., Ren, J., Verma, A., Nettleship, J.E., Owens, R.J., Stuart, D.I., Grimes, J.M. (2006). Lysine methylation as a routine rescue strategy for protein crystallization. *Structure (London, England: 1993)*, 14(11), 1617–22.

Wernimont, A. & Edwards, A. (2009). In situ proteolysis to generate crystals for structure determination: an update. *PloS one*, 4(4), e5094.

Whitaker, J.R. (1963). Determination of Molecular Weights of Proteins by Gel Filtration of Sephadex. *Analytical Chemistry*, 35(12), 1950–1953.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A. & Wilson, K.S. (2011). Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological Crystallography*, 67(Pt 4), 235–42.

Winn, W.J. & Myerowitz, R. (1981). The pathology of the *Legionella pneum*onias. A review of 74 cases and the literature. *Human Pathology*, 12(5): 401–22.

Yu, V.L., Plouffe, J.F., Pastoris, M.C., Stout, J.E., Schousboe, M., Widmer, A., Summersgill, J., File, T., Heath, C.M., Paterson, D.L. & Chereshsky, A. (2002). Distribution of Legionella species and serogroups isolated by culture in patients with sporadic community-acquired legionellosis: an international collaborative survey. *The Journal of infectious diseases*, 186(1), 127–8.

Zuckerman, J.M. (2004). Macrolides and ketolides: azithromycin, clarithromycin, telithromycin. *Infectious disease clinics of North America*, 18(3), 621–49, xi–.

Acknowledgements

It has been a privilege to pursue Ph.D degree in such a supportive and helpful environment in the Department of Molecular and Structural Biology of Helmholtz Centre for Infection Research.

My first debt of gratitude goes to my supervisor, Prof. Dr. Dirk W. Heinz, for giving me the opportunity to work in his group on this project. During my Ph.D study, he provided the vision, the advice essential for my work and completion of my Ph.D dissertation. Furthermore, I am grateful for his support that enabled me to attend a number of international training courses in crystallization and crystallography. These opportunities provided me helpful experiences and knowledge. I am also thankful for his patience and support through difficult time of my study.

I would like to thank Prof. Dr. Dieter Jarn very much for his willingness to be my second supervisor.

I am also grateful that Prof. Dr. Michael Steinert accepted to be the chair of my Ph.D defense examination.

Deeply thanks to Dr. Jörn Kraußе for guiding and helping me dedicatedly throughout my work with immense knowledge, particularly in crystallographic analysis, providing me help and suggestions over the course of the writing of the dissertation.

My great appreciation also goes to Dr. Andrea Scrima for giving me invaluable suggestions in the time of my research and writing of this dissertation. I much appreciate for his detailed proof-readings of my writing and his special patience and support.

I also would like to thank Prof. Dr. Antje Flieger and Katja Kuhle for their collaboration and providing me important information of their functional studies.

I would like to extend my gratitude to Dr. Stefan Schmelz, Dr. Joachim Reichelt and all other members of the Department for offering their precious experience and time to help me solve problems I encountered during my work.

Especially, I wish to extend my heartfelt appreciation to Mrs. Angela Walter. Without her support and care, I might not be able to overcome difficult moments to continue until the end of my Ph.D study.

I would like to thank Mrs. Ute Grummer, Mrs. Christine Bentz, Helmholtz Graduate School for Infection Research for their help in administration works.

And deeply thanks to my family for giving me everlasting and strongest motivation and bringing me to this stage of my life. I could never have completed my Ph.D degree without their love and support. I also would like to send my thanks to my friends for offering me spirit support and wonderful friendship.